

FIGURE 1

ACTGCACCTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA
CCCACGCGTCCGGGCCGGAGCAGCACGGCCGAGGACCTGGAGCTCCGGTGCCTCTTCCC
CAGCGTACCCGCCATGCGCCTGCCCGCCGGGCCGCTGGGGTCTCTGCCGCTTCTGCTG
CTGCTGCCGCCCGCGCCGAGGCCGCCAAGAAGCCGACGCCTGCCACCGGTGCCGGGGCT
GGTGGACAAGTTTAAACAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGCGGGAACA
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCCGCTGCTGGAGATC
CTGGAGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA
GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCAGCTGTCTCGCATGC
CAGGGCGGATCCCAGAGGCCCTGCAGCGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA
GGCGCAGGGCTCTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCAGTACTGCATGG
ACGGTACTTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT
GGACGAGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCTCCCTGCAGCGCTG
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG
GGCTGCACAGGGGAAGGCCAGGAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGAGCA
CGGACAGTGTGCAGATGTGGACGAGTGTCTACTAGCAGAAAAACCTGTGTGAGGAAAAACG
AAAACGTACTACAATACTCCAGGGAGCTACGTCTGTGTGTGCTCTGACGGCTTCGAAGAAACG
GAAGATGCCCTGTGTGCCGCCGCGCAGAGCTGAAGCCACAGAAGGAGAAAGCCGACACAGCT
GCCCTCCCGCGAAGACCTGTAATGTGCGGACTTACCTTTAAATATTTCAGAAGGATGTCC
CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCTCTGCAGTGGACAGCGCGGGAGAGGCTGC
CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGTTGTCTTA
AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG
AGGAAAAAAAAAAAAAAAAAAAAAGGCGGCCGCACTCTAGAGTCGACCTGCAGAAGC
TTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTACAATAAAGCAATAGCA
TCACAAAATTACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTC
ATCAATGTATCTTATCATGTCTGATCGGGAATTAATTCGGCGCAGCACCATTGGCCTGAAAT
AACCTCTGAAAGAGGAACTTGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG
TGTGTCAAGTTAGGGTGTGAAAGTCCCAGGCTCCCAGCAGGCGAAGTATGCAAGCATGC
ATCTCAATTAGTCAGCAACCCAGTTTT

FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop
><MW: 38192, pI: 4.53, NX(S/T): 2
MRLPRRAALGLPLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEKLSKYESSSEIRL
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFPCVKTLKVCCSPGTYPGPDCLACQGGSGRPFCSG
NGHCSGDGSRQGDGSCRCHMGYQGGLCTDCMDGYFSSLRNETHSICTACDESKTCSGLTNRDCGCEVGVWLDE
GACVDVDECAAEPPPCSAQAQCKNANGSYTCEECDSSCVGTGEGPGNCKECISGYAREHGQCADVDECSLAEKT
CVRKNENCYNTPGSYVCVCPDGFETEDACVPPAEAEATEGESPTQLPSREDL

Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343
and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and
313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

FIGURE 3

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCTCCGGGGATCCTCTAGAGATCCCTC
GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCGCCAGCCGCTCTAAACGGGAACA
GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTTGCG
GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGCATGGCCCGGAGGAGCGCCTTC
CCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGC
CGGGCCGCGCAGGAGGAGGCCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA
TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAAATGGCACCTTTTACACATGAT
TTCAGAAAAAGCGCAACAGAGAATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTAC
CTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGAATTCCTGTCCCTTCGCGCTCCCTGGGATA
AAGGCATCATGGCAGATCCAAACCGTCAATGTCCTCTGCTGGGAAACAGTGCCTCACAAGGCA
TCAGTTGTTCAAGTTGGTTTTCCCATGTCTTGGAAAAACAGGATGGGGTGGCAGCATTGGAAGT
GGATGTGATTGTTATGAATTTCTGAAGGCAACACCATTCTCCAACACCTCAAATGCTATCT
TCTTTAAACATGTCAACAAGCTGAGTGCCCGAGGCGGGTGCCGAAATGGAGGCTTTTGTAAAT
GAAAGACGCATCTGCGAGTGTCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTGT
TACCCACGATGTATGAATGGTGGACTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTG
GATTCATGGAGTGAACCTGTGACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACC
TGTTTCTACCTCGGAAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG
CAATGCCCAACCCCTGTGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT
CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGACAT
GGAACCTGCCATGAACCCAAACAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGC
TAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCAGCTCAGGCAGC
ACACGCCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG
TGAACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCA
TGTGTTGAATGTTCAAATAATGTTCACTTAAAGAACTAGGCGCTGAATTTTATTAGCT
TCATTATAAATCACTGAGCTGATATTTACTCTTCTTTAAGTTTCTAAGTACGCTCTGTAG
CATGATGGTATAGATTTTCTGTGTTCACTGCTTTGGGACAGATTTTATATTATGTCAATTGA
TCAGGTTAAATTTTCACTGTGTAGTTGGCAGATATTTCAAATTAATGCACTTTATGGT
GTCTGGGGCAGGGGAACATCAGAAAGGTTAAATGGGCAAAAATGCGTAAGTCACAAGAAT
TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTGAGATATTTAGAT
GTTTGTACATTTTTAAAAATTGCTCTTAATTTTAAACTCTCAATACAAATATATTTTGACC
TTACCATATTTCCAGAGATTCACTATTAATAAAAAAAAAAATTACCTGTGGTAGTGGCATT
AAACAATATAATATTTTAAACACAAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT
TGCTTGAAGCAATATAATATATTGTAAACAAAAACAGCTCTTACCTAATAAACATTTTAT
ACTGTTTGTATGTATAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAATAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAGGCGGCGGCTCTAGAGTCGACCTGCAGAAGCTTGGC
GCCCATGGCCCACTTGTATTATGCAGCTTATAATG

FIGURE 4

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094
><subunit 1 of 1, 379 aa, 0 stop
><MW: 41528, pI: 7.97, NX(S/T): 2
MARRSAFPAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGG
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMSEGNITILQTPQNAIFFKTCQQAECPGGC
RNGGFNCNERRICECPDGFHGHPCHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCTFYPGKCI CPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAERDRP
PESNYIW

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 5

CGGACGCGTGGGCGTCCGGCGGTGCGAGAGCCAGGAGGCGGAGGCGCGGGGCCAGCCTGGG
CCCAGGCCACACCTTACCAGGGGCCAGGAGCCACCA**ATG**TGGCGATGTCCACTGGGGCTAC
TGCTGTTAGCTGCCGTGGCTGGCCACTTGGCTCTGGGTGCCAGCAGGGTCGTGGGCGCCGG
GAGTACGACACCGGTCGACACTGCGGGGCATCCGGGACCGGGAGGCGGTACTGCCAGGA
GCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT
GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC
TGCTCGGCGTGCCACCCCTTTTCCCGCGATCCAAGGATGTATGCATGGAGGTCGTATCTA
TCCAGTCTTGGGAACGTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT
GGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGCTGG
GAACCACAGCGCCCTTCTGGGGCATGACCC**TGGA**TGA**GGG**CATTGCTACCGCCTGGGACCA
TCCGCCATCTTCTCGGTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG
GTGCTTCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCAACTGATTATGAGCCTCTTGA
CCAAGGCAACTGTGTCAGGCTCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT
CAATCCATTCTCTGGGACA CATGACGCCTGTCTGTGCGCCCGAGAACCTGCTGTCTTGTGAC
ACCCACAGCAGCAGGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCTGGTGGTTCTCTGCGTCG
CCGAGGGGTGGTGTCTGACCCTGTACTCCCTTCTCGGGCCCGTGAACGAGACGAGGCTGGCC
CTGCGCCCCCTGTATGATGACAGCCGAGCCATGGGTCTGGGGCAAGCGCCAGGCCACTGCC
CACTGCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACCTCTGTCTACCGCCT
CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA
TGGAGGTGCATGAGGACTTCTTCTTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC
CTTGGGAGGCCAGAGAGATACCGCCGGCATGGGACCCCACTCAGTCAAGATCACAGGATGGGG
AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC
CAGCCTGGGGCGGAGGGGCCACTTCCGCATCGTGC GCGGCGCTCAATGAGTGCACATCGAG
AGCTTCGTGCTGGCGCTCTGGGCGCGCTGGGCATGGAGACATGGGTCATCACTGAGGCTG
CGGCGCACCGCGGGTCCGGCTGGGATCCAGGCTAAGGGCCCGCGGAAGAGGCCCAATG
GGGCGGTGACCCAGCCTCGCCCGACAGAGCCCGGGCGCAGGCGGGCGCCAGGGCGCTAAT
CCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCTGGGAGCCGCGGGCAGGCGAGACTGGCG
GAGCCCCAGACCTCCCACTGGGAGCGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAG
ATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAGCCAGGACACCTCAAGTCTCCAGC
CCCAATACCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCG
TTGCCAGGTTGGAGTGCACTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA
AGTGACCTCCCACTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCAACCACTTCG
TAAATTTTGTATTTTGTGTAAGAGGGGGTCTCACTGTGTTGCCAGGCTGGTTTCGAAC
CCTGGGCTCAAGCGGTCCACCTGCCTCCGCTCCCAAAGTGTGGGATTGACAGGCATGAGCC
ACTGCACCCAGCCCTGTATTCTTATCTTCAGATATTATTTTTTTTCTACTGTTTAAAA
TAAACCAAAAGTATTGATAAAAAAAA

FIGURE 6

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223
><subunit 1 of 1, 164 aa, 1 stop
><MW: 18359, pI: 7.45, NX(S/T): 1
MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQLCCRGRADDC
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMMHGRIYPVLGTYWDNCNR
CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG
```

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,
411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,
269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

FIGURE 7

AGGCTCCTTGGCCCTTTTTCACAGCAAGCTTNTGCNATCCCGATTGTTGTCTCAAATCCA
 ATTCCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
 CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCTGCGTCGCCGAGG
 GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
 CCCCTGTATGATGCACAGCCGAGCCATGGGTGCGGGCAAGCGCCAGGCCACTGCCCACTGC
 CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
 CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
 TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
 AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

$\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^2}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^3}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^4}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^5}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^6}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^7}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^8}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^9}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{10}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{11}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{12}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{13}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{14}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{15}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{16}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{17}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{18}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{19}}$
 $\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i^{20}}$

FIGURE 8

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAAGTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG
CAGCTTGAAGTGCCTGGATGACTCACAGGACTACTACGTGGGCAAGAAGACATCACGTGCT
GTGACACCGAAGTTGTGAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCGGCCAGCTATAGGCTCTGGGGGG
CCCCGTGCAGCCACACTGGGTGTGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG
GCCCAGTGGGAGCCTGTCTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT
CTGCACCCCTGTCCCCACCCCTGACCCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCTCCAACCCCTCTGCTGCTGTTTC
CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCAGGAAGCCTT
CCCTGCCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA
GGGGACAGGCACTCAGGAGGGCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA
GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA
GGGGCCAGGCCTCACATTCGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT
AATAAACACCTGTTGGATAAGCCAAAAAAA

FIGURE 9

MTHRTTTWARTSRAVTPTCATPAGPMPCSRLPPSLRCSLHSACCSGDPASYRLWGAPLQPT
LGVVPQASVPLLTDLAQWEPVLVPEAHPNASLTMVCTVPVPHDPPMALSRTPTRQISSDT
DPPADGPSNPLCCCFHGPAFSTLNPVLRHLFPQEAFFAHPYDLSQVWSVVSPPASRGQALRRAQ

Signal peptide:

amino acids 1-47

N-glycosylation site.

amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.

amino acids 22-26, 76-80

N-myristoylation site.

amino acids 56-60

Amidation site.

amino acids 70-74

100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200

Study	Year	Age	Sex	Sample Size	Prevalence	95% CI
1. [Author et al., 1998]	1998	18-24	F	100	10.0%	5.0-15.0%
2. [Author et al., 2000]	2000	25-34	M	150	15.0%	10.0-20.0%
3. [Author et al., 2002]	2002	35-44	F	200	20.0%	15.0-25.0%
4. [Author et al., 2004]	2004	45-54	M	250	25.0%	20.0-30.0%
5. [Author et al., 2006]	2006	55-64	F	300	30.0%	25.0-35.0%
6. [Author et al., 2008]	2008	65-74	M	350	35.0%	30.0-40.0%
7. [Author et al., 2010]	2010	75-84	F	400	40.0%	35.0-45.0%
8. [Author et al., 2012]	2012	85-94	M	450	45.0%	40.0-50.0%
9. [Author et al., 2014]	2014	95-104	F	500	50.0%	45.0-55.0%
10. [Author et al., 2016]	2016	105-114	M	550	55.0%	50.0-60.0%
11. [Author et al., 2018]	2018	115-124	F	600	60.0%	55.0-65.0%
12. [Author et al., 2020]	2020	125-134	M	650	65.0%	60.0-70.0%
13. [Author et al., 2022]	2022	135-144	F	700	70.0%	65.0-75.0%
14. [Author et al., 2024]	2024	145-154	M	750	75.0%	70.0-80.0%
15. [Author et al., 2026]	2026	155-164	F	800	80.0%	75.0-85.0%
16. [Author et al., 2028]	2028	165-174	M	850	85.0%	80.0-90.0%
17. [Author et al., 2030]	2030	175-184	F	900	90.0%	85.0-95.0%
18. [Author et al., 2032]	2032	185-194	M	950	95.0%	90.0-100.0%
19. [Author et al., 2034]	2034	195-204	F	1000	100.0%	95.0-100.0%
20. [Author et al., 2036]	2036	205-214	M	1050	105.0%	100.0-110.0%
21. [Author et al., 2038]	2038	215-224	F	1100	110.0%	105.0-115.0%
22. [Author et al., 2040]	2040	225-234	M	1150	115.0%	110.0-120.0%
23. [Author et al., 2042]	2042	235-244	F	1200	120.0%	115.0-125.0%
24. [Author et al., 2044]	2044	245-254	M	1250	125.0%	120.0-130.0%
25. [Author et al., 2046]	2046	255-264	F	1300	130.0%	125.0-135.0%
26. [Author et al., 2048]	2048	265-274	M	1350	135.0%	130.0-140.0%
27. [Author et al., 2050]	2050	275-284	F	1400	140.0%	135.0-145.0%
28. [Author et al., 2052]	2052	285-294	M	1450	145.0%	140.0-150.0%
29. [Author et al., 2054]	2054	295-304	F	1500	150.0%	145.0-155.0%
30. [Author et al., 2056]	2056	305-314	M	1550	155.0%	150.0-160.0%
31. [Author et al., 2058]	2058	315-324	F	1600	160.0%	155.0-165.0%
32. [Author et al., 2060]	2060	325-334	M	1650	165.0%	160.0-170.0%
33. [Author et al., 2062]	2062	335-344	F	1700	170.0%	165.0-175.0%
34. [Author et al., 2064]	2064	345-354	M	1750	175.0%	170.0-180.0%
35. [Author et al., 2066]	2066	355-364	F	1800	180.0%	175.0-185.0%
36. [Author et al., 2068]	2068	365-374	M	1850	185.0%	180.0-190.0%
37. [Author et al., 2070]	2070	375-384	F	1900	190.0%	185.0-195.0%
38. [Author et al., 2072]	2072	385-394	M	1950	195.0%	190.0-200.0%
39. [Author et al., 2074]	2074	395-404	F	2000	200.0%	195.0-205.0%
40. [Author et al., 2076]	2076	405-414	M	2050	205.0%	200.0-210.0%
41. [Author et al., 2078]	2078	415-424	F	2100	210.0%	205.0-215.0%
42. [Author et al., 2080]	2080	425-434	M	2150	215.0%	210.0-220.0%
43. [Author et al., 2082]	2082	435-444	F	2200	220.0%	215.0-225.0%
44. [Author et al., 2084]	2084	445-454	M	2250	225.0%	220.0-230.0%
45. [Author et al., 2086]	2					

CCACACGCGTCCGAACCTCTCCAGCGATGGGAGGCCGCCCGCTGCTGCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACC
GACCAGCTGAGCAGGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA
CGTGCAGGTACACGGGCGTTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC
TCATAGTGGAGACGGACACGTTTGGCAGCCGGTTTCGCATCAAAGGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCAGCGGGAAGAGCAAAGACTG
CGTGTTCCAGGAGATCGTGCTGGAGAAACACTATACGGCCTTCCAGAAACGCCCGGCACGAGG
CTGTGTTCAATGGCCCTTACGCGCAGGGGCGGCCCGCCAGGCTTCCCGCAGCCGCGCAGAA
CAGCGCGAGGCCCCACTTCAACGAGCGCCTTACCAAGGCCAGTGCCTTCCCCAACCAACGC
CGAGAAGCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCACCCCGCCGACCAAGCGCACAC
GGCGGCCCCAGCCCCCTCACGTAGCTCTGGGAGGCAGGGGGCAGCAGCCCCGTGGGCGCCTCCC
CACCCCTTTCCCTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCC
GAGGGAGGACCTTGAGGGCCGCGAAGCATCCGAGCCCCAGCTGGGAAGGGGCAGGCCGGTG
CCCCAGGGGCGGCTGGCACAGTGCCTCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT
GAGTGTCAACCCTGATCTCAGGCCACCAGCCTCTGCGGCCTCCAGCCGGGCTCTCTGAAGCC
CGCTGAAAGGTACGCGACTGAAGGCCCTTGAGACAACCGTCTGAGAGTGGCTGTCCTCAAAA
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCAGCCCCCAAACCTCTCTGGCTAGACTGTA
GGAAGGGCTCTTTTGTTGTTTGTGTTTTCAGGAAAAAGAAAGGGAGAGAGAGGAAAAATAG
AGGGTTGTCACTCTTTCATTCTACAGACCAGGCCCTGCACCCCAACCCCAACTCCAGGCC
CGGAATAAAACCACTTTTCTGTC

FIGURE 11

MGAARLLPNLTLCLQLLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRISATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNRGKLIGKPSGKSKDCVFTEIVLENNYTAFQONARHEGWFMFAFTRQGRPRQASRSRQNRQEAHFIKRLYQGQLPFPNHAEKQKQFEFVGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCCTCCCTGTTGAATTTTTTGCACATGGAG
GACAGCAGCAAGAGGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTCTGGACTTCAACAGAACCCTCCAGT
CATTTTGATTTGCTGTTTATTTTTTTTCTTTTTCTTTTCCACCACATTTGTATTTTAT
TTCGGTACTTCAGAAATGGGCCCTACAGACCACAAAGTGGGCCAGCCATGGGGCTTTTTTCCT
GAAGTCTTGGCTTATCATTTCCTGGGGCTCTACTCACAGGTGTCCAACTCCTGGCCTGCC
CTAGTGTGTGCCGCTGCCAGAGGAATTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
CCTCTTGGGATCCCGGAGGGCTAACCGTACTCTACCTCCACAACAACCAATTAATAATGC
TGGATTTCTCGCAGAACTGCACAATGTACAGTCGGTGACACGGTCTACCTGTATGGCAACC
AACTGGACGAATTCCTCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC
AATATTTCAGACCAATTTACGGGCTGCTCTTGCCAGCTCTTGAAGCTTGAAGAGCTGCACCT
GGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCTTCCGGGAGGCTATTAGCC
TCAAATTTGTTGTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCTGTGGAC
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTATATCCGACATGGCCTTCCAGAA
TCTCAGAGCTTGGAGCGTCTTATGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCGG
AGGCCCTTCAGCCCTCCCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC
CACCCTCCTCCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTTCAGGACAACCAAGAT
AAACCAATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
ACAACCAATGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCAACCTGAAGCAGCTC
ACTGCTCGGAATAACCTTTGGTTTTGTGACTGCAGTATTAATGGGTACAGAATGGCTCAA
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCAAGGTCCTGAACAAGTCCGGG
GGATGGCGCTCAGGGAATTAATATGAATCTTTTGCTGTGCCACACAGACCCCGGCTG
CCTCTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACTCTCTAT
TCCAAACCTTAGCAGAAGCTACACGCCTCAACTCTACCACATCGAACTTCCACGATTC
CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTTATC
CATTTTGTAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTCCACGTGATGGCATA
CAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATGTTTCAGGAGCGCATAG
TCAGCGGTGAGAAGCAACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT
TGTTTAGTGCCACTGGATGCTTTAACTACCGCGCGGTAGAAGACACCATTTGTTTCAGAGGC
CACCACCATGCCTCTATCTGAACAACGGCAGCAACACAGCGCTCAGCCATGAGCAGACGA
CGTCCCACAGCATGGGCTCCCTTTCTGCTGCGGGGCTTGATCGGGGGCGCGGTGATATTT
GTGCTGGTGGTCTTGTCTCAGCGTCTTTTGCTGGCATATGCACAAAAGGGGGCGCTACACCTC
CCAGAAGTGGAATAACAACGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
AGGACAACCTCCATCTGAGATGACAGAAAACGATTTTCAGATCGTCTCCTTAAATAACGAT
CACTCCTTAAAGGAGATTTACAGACTGCAGCCATTTACACCCCAAGGGGCATTAATTA
CAGACACTGCCATATCCCCAACCAACATGCGATACTGCAACAGCAGCGTGCAGACCTGGAGC
ACTGCCATACGTGAACAGAGGCCAGCGCTTATCAAGGGCGACAATTAGACTCTTGAGAA
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGAT
TTGTGCAATTTGAATACTCTGTAATTTATACGGGTGTACTATATAATGGGATTTAAAAAAGTG
CTATCTTTTCTATTTCAGTATTAATTACAAACAGTTTTTGTAACCTTTTGCTTTTTTAAATCTT

FIGURE 13

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDNRFVYCNERSLTSVPLGIP
EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI
SRAALQLLKLEELHLDNSISTVGVEDGAFREAISKLFLSKNHLSSVPVGLPVDLQELR
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
PWFCDCSIKWVTEWLKYIPSSLNVRGFMCGQPEQVRGMAVRELNMNLLSCPTTTPGLPLFTP
APBTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPITPDWDGRERVTPPISERIQLSIHFVND
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICALVPL
DAFNRYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSMSGSPFLLAGLIGGAVIFVLVVL
LSVFCWHMHKKGRYTSQKWYNRGRKKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG
DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

FIGURE 15

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPTALLESSCENKRADLVFII
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV
KMRHLSTGIMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD
TGILIFAIGVQVDFTNLKLSIGSEPHEDHVFLVANFSQIETLTSVFPQKKLCTAHMCSTLEHN
CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA
LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCRINCYCALNKPGC
EHECVNMEESYYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSGFLI
NEDLKTCSRVDYCLLDHGCEYSCVNMDRSFACQCPGHHVLRSDGKTCAKLDSALGDHGCE
HSCVSSSEDSFVCQCFEGYILREDGKTCRRKDVCAIDHGCEHICVNSDDSYTCECLEGFRLA
EDGKRCRRKDVCKSTHHGCEHICVNNGNSYICKCSGFLAEDGRRCKCTEGPIDLVFVID
GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNPFNSAKDMKKAVA
HMKYMGKGSMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN
GITMYAVGVGAIEEELQEIASEPTNKHLYAEDFSTMDSEIKLKKGICEALESDSDGRQDS
PAGELPKTVQPTSESEPTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL
EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMALLENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,
401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,
781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,
639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,
540-546, 581-587

FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCAGCTCCGGCCGTCGCGCAGCCTCG
GCACCTGCAGGTCGGTGCGTCCCGCGGCTGGCGCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGCCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGCCAGCTGCAATGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGCGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTGTCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAAACCT
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTTCGTCTTCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCAGGGCCCTGGAGCTGC
AGTGGTTGTGAGGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTAACACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCGGACCCTGCCCTGGGCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCAGGCCTGGTGCAT
TGACCCCCACGCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCCCACCCTCAACCAATATCCCCATCCCTGGTGGGTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGTGCTGCCAGAGTCAAGCTGGCTCTCTGGTATTGATGAC
CCCACCACTCATTGGCTAAGAGATTGGGGTCTCTCCTTCTATAAGGGTCACCTCTAGCAC
AGAGGCCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCTCCACCCACCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTACCAAGAGTGAGGGGCAGAGACTTCAGTCACTGAGTCTCCAGGC
CCCTTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTGAGGCTGGCTGGTTAGGTTTTACTGGGCAGAGGATAGGGAATCTC
TTATTAATAACTACATGAAATATGTGTGTTTTCATTTGCAAATTTAAATAAAGATACATA
TGTTTGTATGAAAAA

FIGURE 17

MISLPGLVTINLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
SQPWVFPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCTRLQGVPHVGANVTLSQCSPRSKPAVQYQ
WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGFGAA
VVAGAVVGTILVGLGLLAGLVLLYHRRGKALEEPANDIKEDIAIAPRTLWPWKSSDTISKNGTL
SSVTSARALRPPHGGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISSIPGGVSSSSGLSR
MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

FIGURE 18

CGCCACCACCTGCGGCCACCGCCAAATGAAACGCCTCCCGCTCCTAGTGGTTTTTTCCACTTTG
TTGAATTGTTTCCTATACTCAAAATGACCAACAGACCTTGTCTCCCAAATGCAAAATGTGA
AATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTCAAA
TTTGTGAAGATGATAATGAATGTGGAATTTAACTCAGTCCCTGTGGCGAAAATGCTAATTGC
ACTAACACAGAAGGAAGTTATTATTGTATGTGTGTAACCTGGCTTCAGATCCAGCAGTAACCA
AGACAGGTTTATCATAATGATGGAAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCAAT
TAGATAATGTCTGTATAGCTGCAAAATATTAATAAACTTTAAACAAAATCAGATCCATAAAA
GAACCTGTGGCTTTGCTCAAGGAAGTCTATAGAAATCTGTGACAGATCTTCCACCAACAGA
TATAATTACATATATAGAAATATTAGCTGAATCATCTTCAATTAAGTTTCAAGAACCAACA
CTATCTCAGCCAAAGGACACCCCTTTCTAACTCAACTCTTACTGAATTTGTAACCAACCGTGAAT
AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC
ACATCTTACAAAACCTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCAGAGCTTCC
AAAAGACCACAGAGTTTGATACAAATTCACCGGATATAGCTCTCAAAGTTTCTTTTTTGAT
TCATATAACATGAAACATATTTCATCCTCATATGAATATGGATGGAGACTACATAAATATATT
TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAAGTGCATTTTTATATTATA
AGAGTTTGGTCTTTTGTCTTTCATCATCTGACAACCTCTTATTGAAACCTCAAAATTTAGAT
AATTCTGAAGAGGAGGAAAGAGTCAATCTTCAAGTAATTCAGTCTCAATGAGCTCAAAACCC
ACCCACATTTATGAATCTGAAAAAATAACATTTACATTAAGTCAATGCAAGAGGTACAGATA
GGTATAGGAGTCTATGTGCATTTTGGAAATTACTACCTGATACCTGAATGCGAGCTGTGCT
TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT
GACACATTTTGCATTTTGATGTCTCTGGTCTTCCATTGGTATTAAGAGATTATAATATTC
TTACAAGGATCACTCAACTAGGAATAAATATTTCAGTGATTTGCTTGGCATATGCATTTTT
ACCTTCTGGTCTTCTCAGTGAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG
TAGCCTATTTCTTGCTGAACCTGTTTTCTTGTGTGGGATCAATACAAATACTAATAAGCTCT
TCTGTTCAATCATTGCCGAGCTGCTACACTACTTCTTTTAGCTGCTTTTGCATGGATGTGC
ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAAACAAGGGATTTTGGCA
CAAGAATTTTTATATCTTTGGCTATCTAAGCCAGCCGTGGTAGTTGGATTTTCCGCGACAC
TAGGATACAGATATTATGGCACAACCAAGTATGTTGGCTTAGCACCAGAAAACAACCTTATT
TGGAGTTTTATAGGACAGCATGCCTAATCATTTCTTGTTAATCTCTTGGCTTTTGGAGTCAT
CATATACAAAGTTTTTGTCTACACTGCGGGTTGAAACCGAAGTTAGTTGCTTTGAGAACA
TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTCTCTTCTCGGCACCACTGGATCTTT
GGGTTTCTCCTGTTTGTGTCAGCATCAGTGGTTACAGCTTACCTCTTACAGTCAGCAATGC
TTTCCAGGGGATGTTTCATTTTTTTATTCTGTGTGTTTTATCTAGAAAGATTCAAGAAGAT
ATTACAGATTGTTCAAAAATGTCCCTGTGTTTGGATGTTTAAAGTAAACATAGAGAATG
GTGGATAATTACAACTGCACAAAAATAAAAATTCGAAGCTGTGGATGACCAATGTATAAAAA
TGACTCATCAATATTCCAATTATTAATACTACTAGACAAAAAGTATTTTAAATCAGTTTTTCT
GTTTATGCTATAGGAACCTGTAGATAATAAGGTAATAATTATGTATCATATAGATATACTATGT
TTTTCTATGTGAAATAGTCTGTGCAAAATAGTATTGCAGATATTTGGAAAGTAATTTGGTTT
CTCAGGAGTGATATCACTGCACCAAGGAAAGATTTCTTTCTAACACGAGAGATATAGAA
TGTCTTGAAGAAACCACTGGCTTGATATTTCTGTGACTGCTGTGCTTTGAACTAGTCC
CCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGAACATAAGAGAATGAAGGGGCAGA
ATATCAAACAGTGAAAGGGGAATGATAAGATGTATTTGAATGAACCTGTTTTCTGTAGAC
TAGCTGGAAGATTTGTGACATAAAATAAGAAATGAAGAAACACATTTTACCATTTGTGAA
TTGTTCTGAACCTTAAATGTCCACTAAAACACTTAGACTTCTGTTTGCTAAATCTGTTTCTT
TTTCTAATATTCTAAAAAAGGTTTACCTCCACAATTGAAAAAAGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 19

MKRLPLL VV FSTLLNCSY TQNC TKT PCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIE NVNANCHLDNV CIAA
NINKTLTKIRSIKEPVALLQEVYRNSVTDLSP TDIITYIEILAESS SLLGYKNNTISAKDTL
SNSTLT EFKVTNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT
NSTDIALKVFFFD SYNMKH IHPHMNDGDYINIFPKRKAAYDSNGNVAVAFLYYK SIGP LLS
SSDNFL LK PQ NYDNSEEEERVISSVIVSMSSNPPTLYELEKITFTLSHRKVTD RYRSLCAF
WNYS PDTMNGSWSSEGC ELYTSNETH TSCRCNHLTHFAILMSSGPSIGIKDY N I LTRITQLG
I I I S L I C I A I C I F T F W F F S E I Q S T R T T I H K N L C C S F L A E L V L F G I N T N T N K L F C S I I A G L
L H Y F F L A A F A W M C I E G I H L Y L I V G V I Y N K G F L H K N F Y I F G Y L S P A V V G F S A A L G Y R Y Y G T
T K V C W L S T E N N F I W S F I G P A C L I I L V N L L A F G V I I Y K V F R H T A G L K P E V S C F E N I R S C A R G A
L A L L F L L G T T W I F G V L H V V H A S V V T A Y L F T V S N A F Q G M F I F L F L C V L S R K I Q E B E Y Y R L F K N V
P C C F G C L R

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636,
648-664

N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181,
188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154,
155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329,
346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394,
434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

[illegible]

TGGAACATATCCTCCCTCATATGAATATGGATGGAGACTACATAAATATATTTCCAAGNG
AAAAGCCGGCATATGGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATATAAGAGTAT
TGGTCCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAATATGATAATTCT
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTAGTCTCAATGAGCTCAAACCCACCCAC
ATTATATGAACTTGAAAAATAACATTTACATTAAAGTCATCGAAAGGTCACAGATAGGTATA
GGAGTCTATGTGGCATTTTGGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAG
GGCTGTGAGCTGCATACTCAAATGAGACCCACACCTCATGCGCTGTAATCACCTGACACA
TTTTGCAATTTTGATGTCTCTGGTCTTCCATGGTATTAAAGATTATAATATCTTTACAA
GGATCAC'CAACTAGGAATAATTATTCTACTGATTGTCTTGCCATATGCATTTTACCTTC
TGTTCTTCAGTGAAATTCAAAGCACCAGGA

FIGURE 21

GCTCCAGCCAAGAACCTCGGGGCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGGCCG
 CTAAGCGAGGCTCTCTCTCCCGCAGATCCGAACGGCTGGGCGGGGTCACCCCGGCTGGGA
 CAAGAAGCCGCGGCTGCTTGCCCGGGCCGCGGGAGGGGGCTGGGGCTGGGGCCGAGGCGG
 GGTGTGAGTGGGTGTGTGCGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
 TGTCTTGGGCACCTACCCGTGGGGCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCCGGAG
 CCGCCGCGCCGTAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCCAACCC
 GGCACCTCACAGCCCCGAGCGCATCCCGGTGCGCGCCAGCCTCCCGCACCCCCATCGCCGG
 AGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGG
 ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTGCGCTTCTCGGACGCGGGGCC
 CCACGTGCACTACGGCTGGGGCGACCCCATCCGCTGCGGCACCTGTACACCTCCGGCCCCC
 ACGGGCTCTCCAGCTGCTTCTGCGCATCCGTGCCGACGCGCTGTGGACTGCGCGCGGGGC
 CAGAGCGCGCACAGTTTGTGGAGATCAAGGCAGTCGCTCTGCGGACCTGGCCATCAAGGG
 CGTGACAGCGTGCAGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT
 ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTATCCGA
 TCCGAGAAGCACGCTCTCCCGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
 CAGAGGCTTCTTCCACTCTCTCATTTCTGCCCATGCTGCCCATGGTCCCAGAGGAGCCTG
 AGGACCTCAGGGGCCACTTGAATCTGACATGTTCTCTCGCCCTGGAGACCGACAGCATG
 GACCCATTGGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCAGCTTTGAGAAGTAACT
 GAGACCATGCCCGGCCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG
 TGCTTCTACAAGAACAGTCTTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAACATCTAGAA
 GTTGTACATATTCAGAGTTTTCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT
 CATAACATTGTAAGCCTGTAGCTTGCCAGCTGCTGCTGGGCCCCCATCTGCTCCCTCGA
 GGTGTGCGGACAAGCTGCTGCACTGTCTCAGTCTGCTTGAATACCTCCATCGATGGGGAAC
 TCACTTCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC
 CCCAGGAGCAGCCAGAAGACAGGCGTAGTTTTAATTTAGGAACAGGTGATCCACTCTGTGA
 AAACAGCAGGTAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
 ACCATTTGCCCTTCCCAAAATCCCTCCAGGCCAGAAGTACTGGAGCAGGCATGGCCCCACAG
 GTTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACCTTGAGAATTCCCC
 CTGAGGCCAGTTCTGTATGGATGCTGTCTGAGAATAACTTGCTGTCCCGGTGTCACTGTC
 TTCCATCTCCAGCCCCACGCGCTCTGCCACCTCACATGCTCCCCATGGATTGGGGCTC
 CCCAGGCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT
 TTGAAGACCCCAAGTCTTGTCAATAACTTGTGTGTGGAAGCAGCGGGGAAGACCTAGAAC
 CCTTCCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA
 TCTCTTATTTTCTTACATTATTTATGCCCCCAATATATTTATGTATGTAAGTGAGGTTTG
 TTTTGATATTAATAATGGAGTTTGGTTGT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

[illegible]

Signal peptide:

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

FIGURE 23

CCCAGAAGTTC AAGGGCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCCTCGACCTCCTCA
GAGCAGCCGGCTGCCGCCCGGGAAGATGCGCGAGGAGAGCCGCCACCGCCTCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCTGCAAAACCCCAA
AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAACTGGGTCGGAGTGTCTCCTTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTCAATATCCG
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCATCTG
AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAAGTGTGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAACTGGAAGT
CTGCAATTTAATACTGTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC
TGTTGGATATCGCAGGTGTCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCGTTTGTCGCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTTCAAAGAAACCTCCTCCAGAAGAGTAATCTTCATCTAA
AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCAGCACTTTGGAAGG
CCGCGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAAATATGGTGAAACCC
CATCTCTACTAAAAATACAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCAGCTGC
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGGTAACAGAGCAAGATTCCATCTCAAAAATAAAATAAAATA
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAATATAGCTTGATATTC

FIGURE 24

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE
WKKLGRSVSFVYYQQTLLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQQNLEED
TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST
NSSYTMNTKTGTLLQFNTVSKLDTGEYSCEARN SVGYRRC PGKRMQVDDLNISGIIAAVVVVA
LVISVCGLGVCYAQRKGYFSKETS FQKSNSSSKATTMS ENVQWLPVIPALWKAAGGSRGQEF

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 130-144, 238-258

N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.

amino acids 226-230

FIGURE 25

GACATCGGAGGTGGGCTAGCACTGAAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG
 AAAGAAGAGGAAGATGTTGGGGCAACATTTATTTTAAACATGCTCCACAGCCCGGACCCCTGGGCAT
 CATGCTGCTATTCTGCAAATACCTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA
 ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA
 TTATATCAITTAAGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTTTTTGGACAATG
 CAATTGTGGCACTGGCACTTATTTCACTGAAGAAAAAATTTGTGGTCTCTATGGCATTATCA
 TTTGACAAATGCAAGCATCTTCTTATCAATCAGCTCCTATTGAACCTTACTAGCACTGACTG
 TGGAACTCTTAAGGGCCATTACATTTTCAAGAAGAAAGCTAAGATGAAGGACATGCCACT
 CCGAATTCATGTGCTACTTGGCCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG
 TGGATTGTCCACGGTTATGTACGTGTGAATCAGGCCCTGGTTTACACCCAGATCCATTTAT
 ATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCAGCCAGATTGCC
 AGCTAACACACAGATTCTTCTCCTACAGACTAACAAATATTGCCAAAATTGAATACTCCACAG
 ACTTTCAGTAAACCTTACTGGCCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT
 ATTAATGTAAAAAAGATGCCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAAACAACTTACTGA
 ACTGCCTGAAAAATGTCTGTCCGAACCTGAGCAACTTACAAGAACTCTATATTAATCACAACT
 TGCTTTCTACAATTTCTAGCTGGAGCCTTTATTGGCCCTACATAATCTTCTTCGACTTCACTCT
 AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT
 TCTGATGATTGGGGAAAAATCCAATTATCAGAATCAAAGACATGAACCTTTAAGCCTCTTATCA
 ATCTTTCGACAGCTGGTTATAGCTGGTATAAACCTCAGAGAAATACCAGATAACGCGTTGGTT
 GGACTGGAACCTTAGAAAGCATCTCTTTTTACGATAACAGGCTTATTAAAGTACCCCATGT
 TGCTCTTCAAAAAGTTGTAAATCTCAAATTTTTGGATCTAAATAAAAAATCCTATTAAATAGAA
 TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAGAGTTGGGGATAAAATAATATGCGCT
 GAGCTGATTTCCATCGATAGTCTTGTCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC
 TACTAACCAACCTGATTTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG
 AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG
 CCAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCTCCGTTG
 GATGAACATGAACAAAACCAATTCGATTTCATGGAGCCAGATTCATGTTTTGCGTGGACC
 CACCTGAATTCCAAGGTCAGAATGTTCCGGCAAGTGCATTTCAAGGACATGATGGAATTTGT
 CTCCTCTTATAGCTCCTGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT
 TTCCTTTCACTGTAGAGCTACTGCAGAACCAAGCCTGAAATCTACTGGATAACACCTTCTG
 GTCAAAAACTCTTGCCATAATACCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA
 GATATAAATGGCGTAACTCCCAAAGAAGGGGGTTTATATACCTGTATAGCAACTAACCTAGT
 TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTCCACAAGATAACAATG
 GCTCTTTGAATATTAATAAAGAGATATTCAGGCCAATTCAGTTTGTGTCTTGGAAAAACA
 AGTTCATAAATCTCAAAATCTAGTGTAAATGGACAGCCTTTGTGCAAGACTGAAAATTTCTCA
 TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC
 CATCAACTGAGTATAAAATTTGTATTGATATTTCCACCATCTATCAGAAAAACAGAAAAAAA
 TGTGTAATGTCAACCAAAAGGTTTGCACCTGATCAAAAAGAGTATGAAAAAGATAATAC
 CACAACACTTATGGCTGTCTTGGAGGCCCTTCTGGGGATTTATGGTGTGATATGTCTTATCA
 GCTGCTCTCTCCCAAAATGAACGTGTGATGGTGGACACAGCTATGTGAGGAATTACTTACAG
 AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA
 AGAAAAAAGTACATCACTGAAAGTAAAGCAACTGTTATAGGTTTACCAACAAATATGTCTCT
 AAAAAACCAAGGAACCACTATCCAAAAATGAAC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 26

MKDMPLRIHVLLGLAITTLVQAVDKKVDPCRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT
FFARLPANTQIILLQTNNAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE
ENKLTLEPEKCLSELSNLQELYINHNLSTISPGAFI GLHNLRLHLNSNRLQMINSKWFDA
LPNLEILMIGENPIIRIKDMNFKPLINLRSIVIAGINL TEIPDNALVGLENLESISFYDNRL
IKVPHVALQKVNLKFLDLNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLPD
LRKIEATNNPRLSYIHPNAFFRLPKLESMLNSNALSALYHGTIESLPNLKEISIHSPNIRC
DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGGQNVVRQVHFRDMMEICLPLIAPESFPSPNINV
EAGSYVSFHCRTAEAPQPEIYWIWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTC
IATNLVGADLKSVMIKVDGSFPQDNNGSLNLIKIRDIQANSVLVSWKASSKILKSSVKWTAFV
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE
YEKNNTTTLMACLGGLLGIIGVICLISCLSPENMCDGGHSYVRNYLQKPTFALGELYPLIN
LWEAGKEKSTSLKVKATVIGLPTNMS

Signal sequence:

amino acids 1-22

Transmembrane domain:

amino acids 633-650

N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,
608-612, 624-628, 625-629

Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.

amino acids 570-579

N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,
491-497, 492-498, 634-640, 702-708

Cell attachment sequence.

amino acids 277-280

FIGURE 27

GCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG
CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CACGCTTGTGTGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCAGCAGTGAATCT
GGTAGACCTGTGGTTAACCCTTCCCTCTCCATGTGTCTCTCTACAAAGTTTGTCTTCTTA
TGATACTGTGCTTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTCTCTCTGGG
GGTTTAAATGTACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCTCTCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTAAAGG
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT
GCCTTCAAAGGAGTAGCTGAAACCTTGCACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACACCCCTGGC
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCAC
AACGTGATCTGTAAAACGTCCTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC
CAACGACGCTGACCTTTGTAACTCCCTAAAAAACTACCGATTATGCCATGCTGGTCAACA
TGTTTGCTGTTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG
GATGCCCGGAGACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATAGTGTCCAACTGACTGTATTGAGAAAAGAAAGAAA
GTAGTTTGCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA
ACTTTGTATTTAGTTTTTTTTGAATTATGCCACTGCTGAACTTTTAAACAACTACAACA
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCGATGGTATATTTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT
AATTTAAAGCAAATAAAGCTTAACTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAACA

FIGURE 28

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMC PKGCLCSSSGGLNVTC SNANLKEIPRDL
PPETVLLYLDNQITSIPNEIFKDLHQLRVNLNLSKNGIEFIDEHAFKGVAETLQTLDSLSDNR
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVI CKTSVLDDEHAGRPFL
NAANDADLCNLPKKT TDYAMLVTMFGWFTMVISYV VYVRQNQEDARRHLEYLKS LPSRQKK
ADEPDDISTVV

Signal sequence:

amino acids 1-33

Transmembrane domain:

amino acids 205-220

N-glycosylation site.

amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation site.

amino acids 37-43, 45-51, 110-116

FIGURE 29

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAGATGCGAGGTGAGCAAGAGGATGCTGGCGGGG
 GGCGTGAGGAGCATGCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCTCCTGCTGGTGCT
 GGGCTCAGTGCTGCAGGCTCGGCCACGGGCTGCCGCCCCGCTGCGAGTGCTCCGCCCAGG
 ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCGAGGGGCATCCCCACCGAG
 ACGCGCTGCTGGACCTAGGCAAGAACCGCATCAAACCGCTCAACCAGGACGAGTTCCGCCAG
 CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGGCG
 CTTTCAACAACTCTTCAACCTTCGGACGCTGGGTCTCCGAGCAACCGCTGAAGCTCATC
 CGCTAGGGCTCTTCACTGGCCTCAGCAACTGACCAAGCAGGACATCAGCGAGAACAAGAT
 CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG
 ACAATGACCTCGTCTACATCTCTCACC CGCCTTCAGCGGCCCTCAACAGCCTGGAGCAGCTG
 ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCACCTGCACGGCT
 CATCGTCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC
 TGTACCGACTCAAGGCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAC
 TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC
 CTACCTGGCCGTCGCCACCTAGTCTATCTCCGCTTCTTCAACCTCTCTACAACCCCATCA
 GCACCAATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCGAGAGATCCAGCTGGTGGGC
 GGGCAGCTGGCCGTGGTGGAGCCCTATGCTTCCCGCGGCCCTCAACTACCTGCGCGTGTCAA
 TGTCTCTGGCAACAGCTGACCACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG
 AGACATCATCTTGGACTCCAACCCGCTGGCCTGCGACTGTCCGCTCTGTGGGTGTTCCGG
 CGCCGCTGGCGGCTCAACTTCAACCGCAGCAGCCACGTGCGCCACGCCCGAGTTTGTCCA
 GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCACTACTTCACTTCCCGCCGCG
 CCGCATCCGGGACCGAAGGCCAGCAGGTGTTTGTGGACGAGGGCCACAGGTGCAGTTT
 GTGTCCGGGCGGATGCGCAGCCCGCCGCCCATCTCTGGCTCTCACC CGGAAAGCACCT
 GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCAGCTGGAGGTGCGCT
 ACGCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACCGGGCGGGCAACGAC
 TCCATGCCCGCCCACTGATGTGCGCAGTACTCGCCGACTGGCCCCATCAGCCCAACAA
 GACCTTCGCTTTTATCTCCAACAGCCGGGCGAGGGAGAGGCCAACAGCACCCCGCCACTG
 TGCCTTTCCCTTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCACTCTTTT
 CTGGCGTCTGCTCTCTTCTGCTGTGCTGCTTTTCTCTGGAGCCGGGGCAAGGGCAACAC
 AAAGCACAACTCGAGATCGAGTATGTGCCCCGAAAGTCGGACGAGGCATCAGCTCCGCCG
 ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGCCCGGGCGGGGGCAGGGACCCCCG
 GCGCGCCGGCAGGGGAAGGGGCTGGTCGCCACCTGCTCACTCTCAGTCTTCCACCTC
 CTCCTACCTTCTACACACGTTCTCTTTCTCCCTCCCGCTCCGTCCCTGCTGCCCCCG
 CCAGCCCTCACCACCTGCCCCCTTCTACCAGGACCTCAGAAGCCAGACCTGGGGACCCCA
 CCTACACAGGGCATTGACAGACTGGAGTTGAAAGCCGACGAAACACACGCGGCAGAGTCA
 ATAATTCAATAAAAAAGTTACGAACCTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTT
 TATGAAAACCTGAAATAATAAAAAAGAGAAAAAACTAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MQVSKRMLAGGVRSMPSPLLACWQPIILLVLGSLSGSATGCPPRCECSAQDRAVLCHRKCF
VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFNNLFLNRTL
GLRSNRLKLIPLGVFTGLSNLTQDISENKIVILLDYMPQDLYNLKSLEVGDNDLVYISHRA
FSGLSNLEQLTLEKCNLTISIPTALSHLHGLIVLRLRHLNINAIIRDYSFKRLYRLKVLEISH
WPYLDTMTPNCLYGLNLTSLSIHNCNLTAVPYLAVRHLVYLRLNLSYNPISTIEGSMLEH
LRLQEIQLVGGLAVVEPYAFRGLNLYLRVLNVSGNQLTLEESVFHSGNLETLILDSNPLA
CDCRLLWVFRRRWRLNFRNQOPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV
FVDEGHTVQFVCRADGDPPPAIILWLSPRKHLVSAKSNGRLTVPDGTLEVRYAQVDNGTYL
CIAANAGGNDSPMAHLHVRYSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPPFPDIKTLI
IATTMGFISFLGVVFLCLVLLFLWSRGKNTKHNIEIYVPRKSDAGISSADAPRKFNMKMI

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 556-578

N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,
492-496, 505-509, 526-530, 542-546

Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,
607-611

Tyrosine kinase phosphorylation site.

amino acids 590-598

N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

FIGURE 31

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAAC
ATCGTAGTCCACCCCTCCCATCCCCAGCCCCGGGGATTGAGGCTCGCCAGCGCCAGCC
AGGGAGCCGGCCGGGAAGCGCGATGGGGGGCCCCAGCCGCCTCGTCTGTCTCTGCTCTGTG
TGTTGCGCTGCTGCTGGGCGCCCCGGCGGGGCCAACCTCTCCAGGACGACAGCCAGCCCTGG
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGTCTCAAGTGCCAAGTGAAGATCA
CGAGGACTCATCCCTGCAATGGTCTAACCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA
GAGCCCTTCGAGATAATCGAATTACAGCTGGTTACCTCTACGCCCCACAGGCTCAGCATCAGC
ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCGCTGT
GCGAATGCCAAGTCCCTCGTCACTGTGCTAGGAATTCCACAGAAGCCCCATCATCACTGGTT
ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAACCTGTGAGTCTTCTGGGAGCAAG
CTTGCAGCCCGGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCAT
ACAGGAAGATCCCAATGGTAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC
GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGAGCTGAC
AGATCCACCTCTCAACGCATTGAAGTTTATACACACCAACTGCGATGATTAGGCCAGACCC
TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGTCGCGGCAATCCAGTCC
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGT
GCCCTGATCTTCCCTTTCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG
CAACATGGGCAGCTACAAGGCCTACTACCCCTCAATGTTAATGACCCAGTCCGGTGCCCT
CCTCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCTGTGTG
CTCATCATGTCTATCTTCTTGGCCACTACTTGATCCGGCACAAGGAACCTACCTGACACA
TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTATCTCTAGAGGCGCCTGCCCACTTCTGTG
GCCCCCAGGGGCCCTGTGGGGACTGCTGGGGCCGTACCAACCCGGACTTGTACAGAGCAA
CCGAGGGCCGCCCTCCCGCTTGTCTCCCAGCCCCACCCACCCCTGTACAGAAATGTCTGC
TTTGGGTGCGGTTTTGTACTCGGTTTGGAAATGGGGAGGGAGGAGGGCGGGGGAGGGAGGG
TTGCCCTCAGCCCTTCCGTGGCTTCTCTGCATTTGGGTTATTATTATTTTGTAACAATCC
CAAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAAGCAAAAAACA
AACAAAAAACA

FIGURE 32

MGAPAAASLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQW
SNPAQQTLYFGEKRALRDNRILQVTSTPHELSSISISNVALADEGEYTCISIFTMPVRTAKSLV
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGK
TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPHPREGQ
KLLHCEGRGNPVPQQYLWEKEGSPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKA
YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLIMLIFLGHYLIRHKGTYLTAEAGSDD
APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,
306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCCTTTTCTTCTCCTTTCTCTGG
 CTTCCGACATTGGAGCACTAAATGAACCTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG
 TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT
 GCTGGAGAGCTCTCTTTGTTTTGCGCGTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA
 TCTGTTCTGCAATGAGATAGAAGGGACCTACACGTAGACTGTGAAAAAAGGGGCTTCACA
 AGCTGCGAGCGTTTCTCGTCCCGGACTTCCCGAGTTTACCATTTATTTCTGTCATGGCAATTC
 CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG
 AAAACAATGGCTTTCGATGAAATCGTTCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG
 CTGCACATCAACAACAAGATCAAGTCTTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA
 TCTGGAATATCTCCAGGCTGATTTTAAATTATACGAGATATAGACCCGGGGGCTTCCAGG
 ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCTTACCTGCCAAC
 GTGTTCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAAACAGGCTGAAAACGCTGCC
 CTATGAGGAGGTCTTGGAGCAATCCCTGGTATTGCGGAGATCCTGTAGAGGATAACCCCTT
 GGGAGTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAAACATTCCCAAGAATGCC
 CTGATCGGCCGAGTGTCTGCGAAGCCCCACCCAGACTGCAGGGTAAAGACCTCAATGAAAC
 CACCGAAGCAGACTTGTCTCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGGCGCCCCCTG
 CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCCTTCAAGACAAATGGGCAAGAG
 GATCATGCCACACCGGGTCTGCTCCAAACGGAGGTACAAGATCCGAGGCACTGGCAGAT
 CAAAAATCAGACCCACGAGCAGCGATAGCGAGGGTAGCTCCAGGAACAAACCTTAGCTAACCA
 GTTTACCTTGCCTTGGGGCTGCAGCTGCGACCACATCCAGGGTGGGTTTAAAGATGAAC
 TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA
 GCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAAATCGCACTTTGTGGATTACAAGA
 ACCTCATTTCTGTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAAACACTTCAAG
 AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTAACCTGGACACGCTGTCCCGGGA
 GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA
 TCCTCCCGGGCACTTTCAATGCCATGCCCAAACCTGAGGATCCTCATTCTCAACAACAACCTG
 CTGAGGTCCCTGCGCTGTGAGCAGTGTTCGCTGGGGTCTCGCTCTCTAAACTCAGCCTGCACAA
 CAATTACTTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACTCCATCATCCAGA
 TAGACCTCCACGGAACCCCTGGGAGTGTCTCTGCACAATTGTGCCTTTCAAGCAGTGGGCA
 GAACGCTTGGGTTCCGAAGTGTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACCTTCTT
 TAGAAAAGGATTTATGTCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
 CGCCCAAGTTAACTTCTGCACAGTAAAAACAGCACTGGGTTGGCGGACCGGGACGCACCTCC
 AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTGGTCCCGGACTGCTGTGGT
 GTTTGTACCTCCGCTTCCACGTTGGTGGGCATGCTCGTGTATTATCCTGAGGAACCGAAAGC
 GGTCCAAAGAGACGAGATGCCAACTCCTCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT
 GACTCTTCTACTGCGACATATGGCCTTACAACGCAGATGGGGCCACAGAGTGTATGACTG
 TGGCTCTCACTCGCTCTCAGACTAAGACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG
 ATACATCCTTCCCCACCGCAGGCACCCCGGGGCTGGAGGGCGGTGACCCAAATCCCGCGG
 CCATCAGCCTGGATGGGCATAAGTAGATAAATACTGTGAGCTCGCACACCCGAAAGGGCCT
 GACCCCTTACTTGTAGCTCCTCCTTGAACAACAGACAGACTGTGGAGAGCTGGGAGAGCGCA
 GCCAGCTCGCTCTTTGCTGAGAGCCCTTTTGACAGAAAGCCAGCAGCACTGCTGGGAAG
 AACTGACAGTGCCTCGCCCTCGGCCCGGGGCTGTGGGGTTGGATGCGCCGGTTCTATAC
 ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTCCCTGTGGATTAG
 CCCCGTGTGGCTCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT
 TGTAAATAAGTAACCTTGTACTTCTGAC

FIGURE 34

MLLWILLLETSLCFAAGNVTGDVCKEIKSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH
LFLHGNSLTRLFPNEFANFNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNKKIKSFRKQ
TFLGLDDLEYLQADFNLRLDIDPGAQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRG
NRLKTLPIYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENI PKNALIGRVVCEAPTRLQ
GKDLNETTEQDLCLPLKNRVDSLPAPPAQEETFAPGPLPTPFKINGQEDHATPGSAPNGGK
IPGNWQIKIRPTAAIATGSSRNKPLANSPLCPGGCSCDHI PGSGLKMN CNRRNVSSSLADLK
KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLDDLRLWLYMDSNY
LDTLSREKFAGLQNLLEYLNVEYNALQLILPGTFNAMPKLRILILNNNLRLSPVDVFAGVSL
SKLSLHNNYFMYLFPVAGVLDQLTSIIQIDLHGNPWECSTIVPFKQWABERLGSEVLMSDLKC
ETPVNFFRDKDFMLLSNDIEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL
VPGLLLVFVTSFTVVGMLVFI LRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG
AHRVYDCGSHSLSD

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 618-638

N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 122-126, 646-650

Casein kinase II phosphorylation site.

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577,
608-612, 657-661, 666-670, 693-697

N-myristoylation site.

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349,
354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 337-348

FIGURE 35

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCCTGACCCAGAACTCAAGGGC
TGCACCGGGCCCTGGCAGCGCTCCGCACACATTTCCTGTGCGCGGCTTAAGGGAACTGTTGGC
CGCTGGGGCCCGCGGGGGATTCTTGGCAGTTGGGGGGTCCGTGCGGAGCGAGGGCGGAGGGG
AAGGGAGGGGGAAACCGGTTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC
AGCTCTGCTCTCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGCGGGGCTCAG
AGAATGAGGGCCGGCGTTTCGCCCTCTGCGCAGGCGCTCTGGCCCGGGCCGGGGCGG
CGGCGAAACCCCACTGCCGACCGTGCTGCTGCTCGGCCCTCGGGGCGCTGTACAGCCTGC
ACCACGCTTACCATGAAGCGGCAGGCGCGCGAGGAGGCGCTGCATCCTGCGAGTTGGGGCGCTC
AGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG
GCCCGGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTGCGACTGGAGCGCAGGCGTTCCTCACT
GCACCCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCTCCGACCCCGCGGCTCTC
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGC
GGTACTCCAGGCCACCGTGGGGTTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC
GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCTGCGCCGCGCCCCGGG
GCCGCTCTAACTTGAGCTATCGCGCGCCCTTCAGCTGCACAGCGCGCTCTGGAAGTTCAG
TCCACCTGGGACCGAGGTGAGTGCCTCTGCCGGGACAGCTCCCGATCTCAGTTACTTGCA
TCGCGGACGAAATCGGCGCTCGCTGGGACAACTCTCGGGCGATGTGTTGTGTCCTGCCCG
GGGAGGTACCTCCGTGCTGGCAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG
CTTTGCGCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA
GTGGGGAAGGACAGCCGACCTTTGGGGGACCGGGTGCCACCAGGCGCCCGCGGCCACT
GCAACAGCCCCGTGCCGAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC
ACCACTTGTCCCTGAACAAGACAATTTCAGTAACATCTATTCTGAGATTCTTCGATGGGGAT
CACAGAGCACGATGTCTACCCCTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCATATCACC
CCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCTCAGGCTTT
CGACTCCTCCTCTGCGCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTGGTGATCT
TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCTCTTCCAGCCA
AGGAAGGAGTCTATGGGCGCGCGGGCTGGAGAGTGATCCTGAGCCGCTGCTTTGGGCTC
CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG
CAGAGGGTGCCCTGCTGGCGGAGTCCCTCTTGGCTCTAGTGATGCATAGGGAAACAGGGGA
CATGGGCACTCTCTGTAACAGTTTTTCACTTTTGATGAACCGGGGAACCAAGAGGAACCTTAC
TTGTGTAACGACAATTTCTGCAGAAATCCCCCTTCCTTAAATTCCTTTACTCCACTGAG
GAGCTAAATCAGAACTGCACACTCCTTCCTGATGATAGAGGAAGTGAAGTGCCCTTTAGGA
TGGTGATACTGGGGACCGGTAGTGTGCGGGAGAGATATTTTCTTATGTTTATTCGGAGAA
TTTGAGAAAGTGATTGAACTTTTCAAGACATTGGAACAAATAGAACACAATATAATTTACA
TTAAAAATAATTTTCAACAAAATGGAAGGAAATGTTCTATGTTGTTTCAGGCTAGGAGTAT
ATTGGTTCGAAATCCAGGGAAAAAATAAAAAATAAAAAATTAAGGATTGTTGAT

FIGURE 36

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS
TVRAGAE LRAVLALLRAGPGPGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCAPAPRGA
ASNL SYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSSEGQPTLGGTGVPTRRPPATA
TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP
SGSVISKFNSTTSSATPQAFDSSSAVVFI FVSTAVVVVLVILTMTVLGLVKLCFHES PSSQPR
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCLDRDRAEGALLAESPLGSSDA

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 189-193, 381-385

Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,
477-483

Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

FIGURE 37

CGGACGCGTGGGATTAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAGTGCCTCCGCCCTGCCGCCGCGTATC
CCCCGGCTACCTGGGCCGCCCCCGCGCGGTGCGCGCTGAGAGGGAGCGCGCGGGCAGCCGA
GCGCCGGTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGCGGTGTGAGCGCGGTGGTGGCGGA
GGGGCGTGTGTGCCGCGCGCGCGCGCTGGGGTGCAAAACCCGAGCGTCTACGCTGCCATGA
GGGGCGCGAACCGCTGGGGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTCTCGCGG
CAGCAGTCCCCAGAGAGACCTGTTTTCACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT
TATTGGCAGTGAAGGTTTTCCTGGAGTGTACCTCCAAATAGCAAAATGTACTTGGAAAATCA
CAGTTCGCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCTATAGACCTCGAGAGTGACAAC
CTGTGCCCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTTGGCCG
CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCGA
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAC
CCCCAAGTGGCCAGACCGGATTACCTTGCAGGAGTCACTTGTGTGTGGCAATTGTAGCCC
CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC
CGATATGATTATGTGGCTGTGTTTAAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA
GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAAATGAACCTCTTATTCACT
TTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTTGGTCACTACATATTTCAGGCCAAAA
AAACTGCCTACAACACAGAACAGCCTGTCAACACCATTCCTGTAAACACGGGTTTAA
ACCCACCGTGGCCTTGTGTCAACAAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT
GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCAACACCATCACTCGCGATGGGAGTTTG
CAGGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTTCAGCAGGCGGG
CAAGAACATGAGTGCCAGGCTGACTGTCTGTGCAAGCAGTGCCTCTCTCAGAAAGAGGTC
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAACAGC
TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG
TTAACAGTGAACTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTC
TCTCAGTAGAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAGATGG
GACTGGTTGACTCTTCATGATGAGGATGATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTT
TGCTGTCTGTCAGAGGAGCAGCTATCTGATTGGAACCTGCCGACTTAGTGCGGTGATAGGA
AGCTAAAGATGTCAAGCGTTGACAGCTTGAAGCGTTTATTTATACATCTCTGTAAAAAGGAT
ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAGATTTTGAAGTGCAATATTTATAGT
GTATTGTGTTTCACTTCAAGCCTTTGCCCTGAGGTGTTACAATCTGTCTTGCGTTTTCTA
AATCAATGCTTAATAAAATATTTTAAAGGAAAAA

FIGURE 38

M R G A N A W A P L C L L L A A T Q L S R Q Q S P E R P V F T C G G I L T G E S G F I G S E G F P G V Y P P N S K C T W K
I T V P E G K V V V L N F R F I D L E S D N L C R Y D F V D V Y N G H A N G Q R I G R F C G T F R P G A L V S S G N K M M V
Q M I S D A N T A G N G F M A M F S A A E P N E R G D Q Y C G G L L D R P S G S F K T P N W P D R D Y P A G V T C V W H I V
A P K N Q L I E L K F E K F D V E R D N Y C R Y D Y V A V F N G G E V N D A R R I G K Y C G D S P P A P I V S E R N E L L I
Q F L S D L S L T A D G F I G H Y I F R P K K L P T T T E Q P V T T T F P V T T G L K P T V A L C Q Q K C R R T G T L E G N
Y C S S D F V L A G T V I T T I T R D G S L H A T V S I I N I Y K E G N L A I Q Q A G K N M S A R L T V V C K Q C P L L R R
G L N Y I I M G V G E D G R G K I M P N S F I M M F K T K N Q K L L D A L K N K Q C

Signal sequence:

amino acids 1-23

N-glycosylation site.

amino acids 355-359

Casein kinase II phosphorylation site.

amino acids 64-68, 142-146, 274-278

Tyrosine kinase phosphorylation site.

amino acids 199-208

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,
305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCACGGCGCCCGCGGGCTGGGGCGGTTCGCTTCTT
CCTTCTCCGTGGCTTACGAGGGTCCCCAGCCTGGGTAAGATGGCCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGGCTCAGCCTCTTCTCAACCTCCAGGACCTATCTGG
CTCCAGCCTCTCCACCTCCCACTCTTCTCCCCGCTCAGCCCCATCCGTGTCTATACCTG
CCGGGGACTGGTTGACAGCTTTAAACAAGGGCTGGAGAGAACCATCCGGGACAACTTTGGAG
GTGGAACACTGCCTGGGAGGAAGAAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCAACAAGCAGCAGGAGGCCCGGACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGTGCCCCGACGGCACCTTCGGGCCCTCTGCG
CTTCCCTGTCTGGGGGAACAGAGAGGGCCTGCGGTGGCTACGGGCAGTGTAAGGAGAAGG
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC
AGTGTGGCCTTGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGT
TTTGGCCCCGTGCCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGAATGCAAGAAGGG
CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGAGCCAAC
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGTCCGAGACTGTGCCAAG
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTGCTGTAAGAAGTGTAGCCCTGGCTATCA
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA
ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGTGTCAGCAGATGTTCTTTGGCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCAGCTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG
ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA
ATCGCGGCCACCACTGTAGGACCTCTCTCCACCCACGCTGCCCCAGAGCTTGGGCTGCCC
TCCTGTCTGGACACTCAGGACAGCTTGGTTTATTTTGTAGAGTGGGGTAAGCACCCCTACCTG
CCTTACAGAGCAGCCAGGTACCCAGGCCCGGGCAGACAAGGCCCTGGGGTAAAGAGTAGC
CCTGAAGGTGGATACCATGAGCTCTTACCTGGCGGGGACTGCGAGGCTTCACAATGTGTGA
ATTTCAAAGTTTTTCTTAATGGTGGCTGTAGAGCTTTGGCCCCGTCTTAGGATTAGGTG
GTCTCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGTGCCAGTTCTGT
TCTGTGTTTACCACATCCCCACCCCATTTGCCACTTATTATTATCTCAGGAAATAAAGA
AAGGTCTTGAAAGTTAAAAAAAAAAAAAAAAAAAAAAAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 40

MAPWPPKGLVPAVLWGLSLFLNLPGPFIWLQPSPPPQSSPPQPHPCHTCRGLVDSFNKGLER
TIRDNFGGGNTAWEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ
QEAPDLFQWLCSDSLKCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDQAG
YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHHLKCVDIDE
CGTEGANCGADQFCVNTGEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLLQQMFFG
IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

Signal sequence:

amino acids 1-29

Transmembrane domain:

amino acids 372-395

N-glycosylation site.

amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 290-294

Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,
326-332, 372-378, 395-401

Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

EGF-like domain cysteine pattern signature.

amino acids 181-193

FIGURE 41

TGAGACCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA
GCACCATGAGCCCCCTGTGGCTCTGCTGGGCACCTCTGGGTGTTGCCCTGGCCAGCCCCGGG
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT
GCCCCCCTGGACAGGGCCGACATGGAGGAGCTGGTTCATCCCCACCCACGTGAGGGCCAGT
ACGTGGCCCTGCTGCAGCGCAGGCCACGGGGACCGCTCCCGCGGAAAGAGGTTTCAGCCAGAGC
TTCCGAGAGGTGGCCGGCAGGTTCTTGGCGTTGGAGGCCAGCACACACCTGCTGGTGTTCGG
CATGGAGCAGCGCTGCCGCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCCGGGCC
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCCTTCGACGTGACCGAGGCCGTGA
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGTACAGGTGTCCGTGCAGAGG
GAGCATCTGGGCCCGCTGGCGTCCGGCGCCACAAGCTGGTCCGCTTTGCCCTCGCAGGGGGC
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCTGGACCTTGGGGACTATGGAG
CTCAGGGCGACTGTGACCTGAAGCACC AATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG
ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAACTGGTGCTGGAGCCCCGGGCTT
CTGGCTTATGAGTGTGTGGGCACCTGCGCGCAGCCCCGGAGGCCCTGGCCTTCAAGTGGC
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCAACATGAGGGTGCAGAA
GTGCAGCTGTGCCTCGGATGGTGCCTCGTGCCAAGGAGGCTCCAGCCATAGGCGCCTAGTG
TAGCCATCGAGGACTTGACTTGTGTGTGTTTCTGAAGTGTTCGAGGGTACCAGGAGAGCTG
CGGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT
CCTTGACAAGTTACCTCACCTAATTTTGTCTCTCAGGAATGAGAATCTTTGGCCACTGGA
GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCCANTGTGTTCATGTGTTTACTTGTCTGTAC
TGGATCTGGGCTAAAGTCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT
AAAA

FIGURE 42

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQKVEPTLDRADMEELVIPTHVRAQYV
ALLQRSHGDRSRGKRFSQSFRVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP
VPKAALHRHGRLSPRSARARVTVEWLVRVDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF
WQQLSRPRQPLLLQVSVQREHLGLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCTCRQPEALAFKWP
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSASD GALVPRRLQP

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

FIGURE 43

GTCTGTTCCAGGAGTCCTTCGGCGGCTGTTGTGTGTCAGTGGCCTGATCGCGATGCGGGACAAA
GGCGCAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCTGTGCCTACTCGGGCTTTTCTTCTCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGTCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACCTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGTCCTCTGAGGAAGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGTGACATGCTCAGAAACAAGATGGTTCCTCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCTATGTCTCT
GAATCCCAACAACAGGAGAGCTGGTCTTTGATCCCCTGTGAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACCGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGTTCATCGTGGCAGCCGCTTGTGAACCTGATTTCTCTGGGAAT
CTTGGTTTTTGGCATCTGGTTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTACAGCCAGCCTAGTCCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCTCTGGTGTGAGCCTGGTTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGTACCGACTCTGGCCCTGATGTCTGTAGTTTCAAGGATGCCTTATTGTCTTCTC
TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTCCTACCCTGCTGAGTGGCTGGAACCTTGTTTTAAA
GTGTTTTATCCCAATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGTCGAGGAATCTGCACTCAACTGCCACCTGGC
TGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
TGATGACACTGGGGTCTTCCATCTCTGGGGCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCTCTCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAATGGGAGCTCTTGTGTGGAGAGCATAGTAAATTTTCAGAGAACCTGAAGCCAAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAAGAAAACCTGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA
ACATGGAGAAAACCTACTGGAATAACAAGTTAGCCAGGCATGGTGTGTCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

FIGURE 44

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGDTRTLVCYNNKITASYEDRVTFLLPTGITPKSVTREDTCTYTTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPEYTWFKDGI VMP TNP KSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGT P MTSNAVRMEAVERNVGVIVA AVLVT LI
LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

FIGURE 45

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGA
GCGTGGCGAAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTCGGCCTCGGACTAGGCCCT
GGAGGCCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGAGGCCCCAGCTCAG
GCTCGTGCCCAACCACCAAGTTCCAGTGCCGACCACTGGCTTATGCGTGCCCTCACCTGG
CGCTGCGACAGGGACTTGACTGTCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC
ATGTACCCAGAAAGGGCAATGCCACCGCCCCCTGGCCTCCCTGCCCTGCACCGGCGTCA
GTGACTGCTCTGGGGGAACTGACAAGAACTGCGCAACTGCAGCCGCTGGCCTGCCCTAGCA
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGEGCCA
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG
GGGATGCCACAACCATGGGGCCCCCTGTGACCCCTGGAGAGTGTACCTCTCTCAGGAATGCC
ACAACCATGGGGCCCCCTGTGACCCCTGGAGAGTGTCCCCTCTGTGCGGAATGCCACATCCTC
CTCTGCCGAGACCACTGTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC
TCAGTGCAAGCCTGGTCAACGCCACCCTCCTCTTTGTCTGGCTCCGAGCCAGGAGCGC
CTCCGCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC
CTCGCTGCCCTGAGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGACACA
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC
TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACTGCCACAGCCAGAACTGAG
GGGCTGGCCCCAGGCAGCTCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

FIGURE 46

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR
TSGLCVPLTWRCRDLDCSDGSDEEECRIFPCTQKGQCFFFFPGLPCPCTGVSDCSGGTDKKL
RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEIILPEGDATTMGPPVT
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLL
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

Signal sequence:

amino acids 1-30

Transmembrane domain:

amino acids 230-246

N-glycosylation site.

amino acids 126-130, 195-199, 213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,
224-230, 230-236, 263-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

Leucine zipper pattern.

amino acids 17-39

FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGGCGGCAGCAGAGGTGCGGCACAGATGCGG
GTTAGACTGGCGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG
GAGATGGATTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTCTACTGCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTAATGGAACCTTAGGCTGGATCCCAAGTGA
TAATTCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCTATA
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCTATGAAGGATTCAAGATCCGG
TACCCCGACCTACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCC
CATCTGTCAAGGCTGCGCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT
GATGGGTCTGCGTATCTTTGAGTGCTTACAAAACCTTATCTGGTCTGCAGCCACGCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTACGGAGATTTCTGCT
GCCACCCGCGGCTTGTGAGCGCTACAACCACGGAACTGTGGTGAGTTTTACTGCGATCCT
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCTTTC
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCAGCACCCATGAGACCCTCCTGA
CCACGTGGAAGATTGTGGCGTTACGGCAACCAGTGTGCTGCTGGTGTCTGCTCGTCATC
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCACTTTCCCCCAGGGGGCCTCCCCGGAG
TTCCAGCAGTGACCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG
ACGAAGCTGTGAGTGGCGGCTTGAGTGCCCTTAGGCCCGGGTACATGGCCTCTGTGGGCCAG
GGCTGCCCTTACCCTGGACGACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGA
CACAGGCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA
GTCTGTATTACCTCCAGGTGCCAAGAGAGCACCACCCTGCTTCGGACAAACCTTGACATA
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCAGGCATCCATATGCCCCTGGGT
GTTGTTCTTAAGAAACTGATTGATTAAAAAATTTCCCAAAGTGCTCTGAAGTGTCTCTTCAA
ATACATGTTGATCTGTGGAGTTGATTCTTTCTCTCTGGTTTTAGACAAATGTAAACAA
AGCTCTGATCCTTAAATTGCTATGCTGATAGAGTGGTGAGGCTGGAAGCTTGATCAAGTC
CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAAAAGNAAAAAA

FIGURE 48

MYHGMNPSNGDGFLEQQQQQQPQSPQRL LAVILWFQLALCFGPAQLTG GFDLQVCADPGI
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGLGWIPSDNSICVQEDCRI
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGLRPLAS
SNGYVNISELQTSFFVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPRCLALEAQVCPLP
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT
WPSTHETILLTTWKIVAFATSVLLVLLLVILARMFQTKFKAHFPPRGPPRSSSDPDFVVVD
GVPVMLPSYDEAVSGGLSALGPGYMASVQGCP LPVDDQSPPAYPGSGD TDTGPGSEETCDS
VSGSSELLQSLYSPRCQESTHPASDNPDIIASTAEVASTSPGIHHAHWLFLRN

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 325-344

N-glycosylation site.

amino acids 104-108, 134-138, 192-196

Casein kinase II phosphorylation site.

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,
364-368, 380-384, 467-471, 468-472

N-myristoylation site.

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,
478-484

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 394-405

FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCGCCTCCCGTGC GGTCCGTCGGTGGCCTAGAGA
TGCTGCTGCCGCGGTTGCAGTTGTGCGGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT
AGCGCCCGAGTGTGCGGGGGCGCACCCGAGTCGGGCCATGAGGCCGGGAACCGCGCTACAGG
CCGTGCTGTGGCCGTGCTGCTGTTGGGGCTGCGGGCCGCGACGGGTCGCTGCTGAGTGCC
TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGAGGGACACAGAGGCCCTTGTTA
TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACCTTTGAGGAAGCCAAAGAAGCCT
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAACTGATAGAA
AAGTTCAATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA
GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT
CACAATTTAGGAACTGGTATGTGGATGAGCCGCTCTGCGGCAGCGAGGTCTGCGTGGTCATG
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGGAATGATGA
CCGTTGCAACATGAAGAAACAATTTCAATTTGCAAATATTCGTAGAGAAACGAGCAGTTCCTT
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG
GAAGAAGATGCCAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT
CCTAATCCCCAGCATTTCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTGGGTTT
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACACCATC
TGGCCCTCTCCTCACCAGGGAAACAGCCCGACCTAGAGGTCTACAATGTCATAAGAAAAACA
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCAATCCGAGTGTGTT
CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA
GAAAGTGGGTTTGTGACTCTGCTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA
GTTCTCCCAGACCAATGAGGAGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG
GTTATTAGGACATATAAAAACTGAAACTGACAACAATGGAAGAAAGAAATGATAAGCAAAATC
CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCTGT
GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT
CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCAGGTCTGGCACATAGTA
GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTTAAAGGACAGAGCTTTACCTG
GCAGTGATAAAGATGGGCTGTGGAGCTTGGAAAACCACTCTGTTTTCTTCTCTATACAG
CAGCATATATTATCATACAGACAGAAAATCCAGAACTTTTCAAAGCCACATATGGTAGCACAG
GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTCAAAGAATAAAATCAAATAAAGA
GCAGGAAAAAAA

FIGURE 50

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGD FWIGLR RREBKQSNSTACQDL
YAWTDGSI SQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGETELTTPVLPEETQEEDAKKTFKESREAA NLAYILI PSIPLLLLLV
VTVVVCVWVICRKRKREQPD PSTKKQHTIWPSPHQGN SPDL EYVNVIRKQSEADLAETRPDL
KNISFRVCSGEATPD DMSCDYDNMAVNPSES GFTLVSVESG FVTNDIYEFSPDQMGRSKES
GWVENEIYGY

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,
299-303, 306-310, 323-327

N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

FIGURE 51

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC
ATCCGCAGGTTCCCGGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT
GTTTGCCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT
GTTCAGCATGCGCTTGTGGACCCAGTGGGCGTCTGACCTCGTGGCGTACTGCCTGCACC
AGCGGGCGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC
CTGTGTAAGTTGAAATGGTGCAGGTCTGTTTCGACACGGGGCTCGGAGTCTCTCAAGCC
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAAACCCCACTATTAGAGGTCCACCCCAAACTC
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCTTACGACTCT
CAATACCATGAGACCACCCTGAAGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT
GCAGCAAATGTTTGCCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC
TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTCGTTCCACTAACATTTTTCGGAATCTG
GAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTCCAGTGTGAGAAAGAACCCATCATCAT
CCACACTGATGAAGCAGATTGAGAACTCTGTATCCCACTACCAAGCTGTGAGGCTGA
GGCAGAGAACAGAGGCCGGAGGCAGACTGCCTCTTACAGCCAGGAATCTCAGAGGATTG
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTG
CACGGATGATGAACAGAGAGCTGTGGACACATCCTGTACATACTGCCAAGGAAGACAGG
GAAAGTCTTCAGATGGCAGTAGGCCATTCCTCCACATCTAGAGAGCAACCTGCTGAAAGC
CATGGACTCTGCCACTGCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG
TGACCTTCATACCGCTCTTAATGACCCCTGGGGATTTTGTACCACAAATGGCCACCGTTTGCT
GTTGACCTGACCATGGAACCTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT
TCTTGAATGCCATGTTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGGAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT
TTAAAATAAAGTGCCCTTTATACAATG

FIGURE 52

MITGVFSMRLWTFVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLMVQVVFRHGAR
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL
TKVGMQQMFPALGERLARKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE
GPILIHTEADSEVLYPNYQSCWSLRQRTGRRRQTASLQPGISEDLKVKVDRMGIDSSDKVD
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFPHILES
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLMLTLGIFDHKWPPFAVDLTMELYQHLESKEW
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVYTLSPKYHALCSQTQVMEVGNNE

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 218-222

Casein kinase II phosphorylation site.

amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.

amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.

amino acids 50-65

FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT
TAAATTTAGCTCATCACCTTCACCTGCCTTGGTCATGGCTCTGCTATTCTCCTTGATCCTT
GCCATTTGCACACAGACCTGGATTCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG
ACGGCTGGGACATTAAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG
ATTGTTACATGATGAAGATGCTGGGGCATCGTGTGAGAAGCCAGAGAGCTCTTTCTCCCA
GTCCAGAGGGTGTGAGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA
GCACCAGAACAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCCGCAAGGTGG
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCTTCA
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACAGTGGGTG
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTGCTTGGCTCAGGGGAGGAG
CAGTCCCTGGAGCAGTGCACGACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA
TGTGGCTGTCTCTGCTCAGTGTAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGATGAACACTGATCT
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGGCCCTGATTCTCAGGCCCTTCAGAGTTGG
ATCAGAAGCTTACAACATCAGGCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA
CCACCTTTCTATGTCTCCACATGTCACACAGCAGATTTCCAGCCTCCATAATTGTGTGTAT
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACATA
CACCATTGTCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTTTAAAGATAAAATTTCTGAATTGGTTATGGGGTTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATCTGGTAACCTTTATTTACAATAATAAGATAGCAC
TATGTGTTCAA

FIGURE 54

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCGRVEVEQKGQWGTVCDDGWDIKDVAVLC
RELCCGAASGTPSGILYEPFAEKEQKVLIQSVSCTGTEDTLAQCEQEVEYDCSHDEADAGASC
ENPESSEFSPVPBGVRLADGPGHCKGRVEVKHQNQWYTVCTGWSLRAAKVVCRLGCGRVLA
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPGWKNTCNHDEDTWVECEDPPDLRLVG
GDNLCSGRLEVLHKGWVGSVCDDNWGEKEDQVVCQLGCGKSLSPSFRDRKCYGPGVGRIWL
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

Signal sequence:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,
267-271, 294-298, 316-320, 336-340

N-myristoylation site.

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,
180-186, 263-269, 286-292

Amidation site.

amino acids 196-200

Speract receptor repeated domain signature.

amino acids 29-67, 249-287

bioRxiv preprint doi: <https://doi.org/10.1101/2017.07.11.166666>; this version posted July 11, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCTCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCGGGCTACCAGGAAGAGTCTGCCGAAG
GTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCTGCTGTTCCGGTGCCTGGG
CGTCTTCGGCCTCTTCGGGTGCTGCAGTGGGTGCGCGGAAGGCTTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAGAATGTGCAAAAGTCTTCTATGCT
GCGGTGCTAAACTGGTGCTCTGTGGCCGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCGCCAAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC
GACATACTTGTCACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGA
TGTGGACAAGAGGGTCATGGAGACAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCCTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCAGGCTTTCTTTGA
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA
TCCACACCAACCTCTCTGTAAATGCCATCACC GCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACCACAGCCCGAGGCCGAAGCCCTGTGGAGGTGGCCAGGATGTTCTTGCTGCTGTGGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTATCTTCGAA
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTTTAGGCTTG
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG
AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA
AGCTTCTTCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT
AAAAACTAGAAATAAACATCTCAAACAGTAAAAAAAAAAAAAAAAAGGGCGGCCGCACTCTAG
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTATTATGCAGCTTATAATGGTTAC

FIGURE 56

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA
KLVLCGRNGGALBELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL
VNNAGISYRGTIMDTTVDVKRVMETNYFGPVALTKALLPSMIKRRQGHIVAIISSIQGKMSI
PFRSAYAASKHATQAFDFCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSGRYGVMDTTT
AQGRSPVEVAQDVLAAVGGKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 104-120, 278-292

N-glycosylation site.

amino acids 228-232

Glycosaminoglycan attachment site.

amino acids 47-51

Casein kinase II phosphorylation site.

amino acids 135-139, 139-143, 253-257

Tyrosine kinase phosphorylation site.

amino acids 145-153, 146-153

N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.

amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

[illegible][illegible]

FIGURE 58

MKFLLDILLLLPLLVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK
LVLWDINKHGLEETAACKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV
YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHI VTVASAAGHVSV PFLLAYC
SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVNRLMH
GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMQ

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34, 283-287

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

FIGURE 59

CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCGCGGCTC
 AGGGAGGAGCACCGACTGCGCCGACCCCTGAGAGATGGTGGTGCCATGTGGAAGGTGATTG
 TTTCGCTGGTCTGTGATGCCTGGCCCCGTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT
 GTTTCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC
 TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCAGGACTGAACATGA
 AGAGTTATGCCGGCTTCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTCTGGTTC
 TTCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG
 AGGTTCCATCCATTGTTGACTCTTTGTGGAACATGGGCCTTATGTTGTGACAAAGTAACATGA
 CTTGCGTGACAGACACTTCCCCTGGACCAACGCTCTCCATGCTTTACATTGACAATCCA
 GTGGGCACAGGCTTCAGTTTACTGATGATACCCAGGATATGCAGTCAATGAGGACGATGT
 AGCACGGGATTTATACAGTGCACTAATTCAGTTTTCAGATATTTCTGAAATATAAAAAATA
 ATGACTTTTATGTCACTGGGGAGTCTTATGACGGGAAATATGTGCCAGCCATTGCACACCTC
 ATCCATTCCCTCAACCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA
 TGGATATTCTGATCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT
 TGTGGATGAGAAGCAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC
 AGGAAGCAGAACTGGTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC
 AAGTGATCCTTCTTACTTCCAGAAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGT
 GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTGTCACTCCCAGAGGTGAGACAA
 GCCATCCCGTGGGGAATCAGACTTTAATGATGGAACATAGTTGAAAAGTACTTGGCAGA
 AGATACAGTACAGTCAAGTAAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA
 TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC
 ATGGACTGGAAGGATCCAGGAATACAAGAAGGCAGAAAAAAAGTTTGAAGATCTTTAA
 ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC
 GAGGTGGAGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAAATCGA
 TTCATTTATGGAAGGATGGGATCCTTATGTTGGAATAACTACCTCCCAAAGAGAACAT
 CAGAGGTTTTCTATTGCTGAAAGAAAATCGTAAAAACAGAAAATGTCTAGGAATAAAAAAA
 TTATCTTTTATATCTGCAAGATTTTTTTCATCAATAAAAAATATCCTTGAAACAAGTGAGC
 TTTGTTTTTGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA
 TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA
 AATTTTAGGGTCTGAAATAGGAAGTTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG
 TAACAAACAAAGCTGTAACATCTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT
 GTTTGGAATATTATTGGATAAGAATAGCTCAATTATCCCAAATAAATGGATGAAGCTATAA
 TAGTTTTGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA
 AAAATATTATATATAAAGTAAAAAATAA

FIGURE 60

MVGAMWKVIVSLVLLMPGCDGLFRSLYRSVSMPPKGDGQPLFLTPTYIEAGKIQKGRELSL
VGPPFGLNMKSYAGFLTIVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH
GPYVVTSMNTLDRDFPWTTLTSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF
FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSNLNPVREVKNLNGIATGDGYSDPESIIGG
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGLTSDPSYFQNVGT
CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT
EIMNNYKVLIIYNGQLDIIVAAALTEERSLMGMDWKGSGEYKKAEEKVWVKIFKSDSEVAGYIRQ
AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYKGKWDPPYVG

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,
353-357, 424-428

Tyrosine kinase phosphorylation site.

amino acids 423-432

N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

FIGURE 61

CGAGGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTGGCTACAAACAT
 TTTTCCCTTTCCTAACAAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT
 GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGTGACAGCTCTCTTG
 CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG
 TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC
 CTGGGGGAGGGCCCTGCCTAACCAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT
 AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGG
 TGAACTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAAACGATGGCTTAAGGGGCCAGAAA
 TAGAGATGCTTTGTAAAATAAAAAATTTAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA
 GACCAAAATAGATAACAGGATTCCCTGAACATTCTTAAGAGGGAGAAAAGTATGTTAAAAATA
 GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCCTGGGTC
 AGGCCAGCCTCTTTGCTCCTCCCGAAATTATTTTGGTCTGACCACCTCTGCCTTGTGTTTT
 GCAGAATCATGTGAGGGCCAAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT
 CCTCACCGCGCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG
 ACAGCCGCTCTGTGGTCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCAGCAGCCGGC
 ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT
 CCACCAAGGGACGGGGCCCTCTATGTGGGGCCATCAACCGGGCTATAAGCTGACAGGCA
 ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG
 CCCCCTCATCGTGCAGCCTTGACGCGAAGTGCTCACCCCTCACCAATGTCAACAAGCTGCT
 CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTTACCAGGGGTCTGCA
 AGCTGCTGCGGTGGATGACCTCTTCATCCTGGTGGAGCCATCCCAAGAAGGAGCACTAC
 CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA
 TGGCAAGCTCTTCATGGCAGGCTGTGGATGGGAAGCAGGATTACTTCCGACCCCTGTCCA
 GCCGGAAGCTGCCCGAGACCCTGAGTCTCAGCCATGCTGCATATGAGCTACACAGCGAT
 TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCCTGGCCCTGGTCTCCCACTTTGACAT
 CTTCTACATCTACGGCTTTGTCTAGTGGGGCTTTGTCTACTTTCTACTGTCCAGCCCGAGA
 CCCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTACGCATCGTGGG
 CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC
 CGGGGTGGAATACCGCCTCTGCAAGGCTGCTTACCTGGCCAAAGCCTGGGGACTCACTGGCCC
 AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCAAAGGGCAGAAG
 CAGTATCACCACCCGCCGATGACTCTGCCCTGTGTGCCCTTCCCTATCCGGCCATCAACTT
 GCAGATCAAGGAGCGCTCTGACTCTGTACAGGGCGAGGGCAACCTGGAGCTCAACTGGC
 TGCTGGGGAAGGACGCTCCAGTGCACGAAGGCCTGTCCCCATCGATGATAACTTCTGTGGA
 CTGGACATCAACCAGCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCTGTACACCA
 CAGCAGGACCGCATGACCTCTGTGGCTCTTACGTTTACAACGGCTAAGACGCTGGTTTTTG
 TGGGACTAAGAGTGGAAGCTGAAAAAGGTAAGAGTCTATGAGTTGAGATGCTCCAATGCC
 ATTCACTCTCAGCAAAGAGTCCCTCTTGAAGGTAGCTATTGGTGAGATTTAACTATAG
 GCAACTTTATTTCTTGGGGAACAAAAGGTGAATGGGAGGTAAGAAGGGGTTAAATTTGTG
 ACTTAGCTTCTAGCTACTTCCCTCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA
 TTTCATATTTCCCAAACCTTTAAGAAAAAAGCTTTAAGAGGTACATCTGCAAAAGCAGA

FIGURE 62

MGTLGQASLFAPPNGYFWSHDHSAFCFAESCEGQPGKVEQMSTHRSRLTAAPLSMEQRQPWP
RALEVDSRSVLLSVVWVLLAPPAAGMPQFSTFHSNDRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPPPLIVQPCSEVLTLTNVVKLLIIDYSENRLACGSL
YQGVCKLLRLDDLFIIVPEPHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFITAVDGKQDY
FPTLSSRKLPDPPESSAMLDYELHSDVFVSSLIKIPSDTLALVSHFDIFYIYGFASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAAQAFNITSQDDVLFATFSKGQKVHHPDDSAACAPPIRAINLQIKERLQSCYQGEEN
LELNWLLGKDVQCTKAPVPIDDFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTKSGKLLKVRVYEFRCNSNAIHLLSKESLLEGSYWWRFNRYRQLYFLGEQR

Signal sequence:

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,
384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

FIGURE 63

AGGCTCCCGCGCGCGGCTGAGTCCGGAAGCTGGAGTGGGAACCCGGGTCCCCCGGCTTAGAGAACACGCGAGTACCA
 CGTGGAGCCTCCGCGGAGGCGCGGCCGACGCTGGGACTCCTGCTGCTGTGCTCTTGGGCTTCTGTGTGCTCC
 GCAGGCTGAGCTGGAGCACCTGTGTCTCTGCGGCTCCGCCATCGACAGCTGGGCTGCAGGCCAAGGCTGGGA
 ACTTCATGCTGGAGGATTCCACTTCTGGATCTTGGGGGCTCCATCCACTATTCCGTGTGCCCGAGGGAGTACT
 GGAAGGACCGCTGCTGAAGATGAAGGCTGTGGCTTGAACACCTCACCACTATGTTCCGTGGAACCTTGCAATG
 AGCCAGAAAGAGGCAAAATTGACTTCTCTGGGAACCTGGACCTGGAGGCTTCTGCTGCTGATGGCCGAGAGATCG
 GCGCTGTGGGTGATTCTGCGTCCAGGCCCTTACATCTGCAGTGAGATGGACCTCGGGGGCTTGCCCACTGGCTTAC
 TCCAAGACCTGGCATGAGGCTGAGGACAACTTACAAGGGCTTCAACGAAGCAGTGGACCTTATTATTGACCACT
 TGATGTCCAGGGTGGTCCACTCCAGTACAAGCGCTGGGGGACCTATCATTGGCTGCAGGTGGAGAATGAATATG
 GTTCTCTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCATCTGGAGGACCGTGGCATTGTGGAACTGCT
 TCCGTACTTCAGACAACAAGGATGGCTGAGCAAGGGGATTGTCCAGGAGTCTTGGCCACCATTCAACTTTCAGT
 CAACACACGAGCTGCAGCTACTGACCACTTTCTCTTCAAGCTCCAGGGGACTCAGCCCAAGATGGTGTGAGGT
 ACTGGACGGGGTGGTTTGACTCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCTGT
 CTGCCATTGTGGACGCGGCTCTCCATCAACCTCTACATGTTCCACGGAGGCCAACCTTTGGCTTCATGAATG
 GAGCCATGCACTTCCATGACTACAAGTCAAGTGTACCAAGTATGACTATGATGCTGTGAGGTTTTGAAAACCTGT
 ATTACACGGCCAAAGTACATGAAGCTTGCAGACTTCTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCCACTG
 ACCTTCTTCCCAAGATGCGGTATGAGCCCTTAAACGCGAGTCTGTACTGTCTCTGTGGGACGCTTCAAGTACC
 TGGGGGAGCCAAATCAAGTCTGAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGCGAAATGGACAGTCTCT
 TCGGGTACATTCTCTATGAGACCAAGTCACTCCTGCTCTGGCATCCTCAGTGGCCAGCTGCATGATCGGGGGCAGG
 TGTGTTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCTGATCCAGGGTT
 ACACCGTCTGAGGATTCTGTGGAGAATCGTGGGCGAGTCAACTTGGGGGAGAATATTGATGACCCAGCGCAAG
 GCTTATTTCGAAATCTCTATCTGAATGATTCAACCTTGAAAACTTCAGAACTATAGCTGGATATGAAGAAGA
 GCITCTTGGAGGTTTCCGCTTGAGCAAAATCGTGGGCGAGTCAACTTGGGGGAGAATATTGATGACCCAGCGCAAG
 GTAGCTTGTCCATCAGCTCAACGCTTGTGACACTTCTGAAGCTGGAGGGCTGGGAGAAGGGGGTGTATTCA
 TCAATGGCCGAACCTTGACGCTTACTGGAACATTGGACCCAGAGAAGCGCTTACCTTCCAGGTCCTGTGTTGA
 GCAGCGAATCAACAGGCTACGTTTTGAGGAGACGATGGCGGGCTGCATTACAGTTTCAAGGAAACCCCC
 ACCTGGGCGAGAACCTAGTACTTAAGTGAAGCGGTGGCACCCCTCTGCTGGTGGCAGTGGGAGACTGCGCGCTC
 CTCTTGACTTGAAGCTGTGGGCTGTGCCCAACCCCTCACTGCAAAAGCATCTCTTAAAGTACCAACCTCAGG
 ACTGGGGCTACAGTCTGCCCTGTCTCAGCTCAAAACCTTAAAGCTTCAAGGAGGAGTGGGATGGCTCTGGGCT
 TGGCTTGTGATGATGCTTCTCAAGCCCTGCTCTGTGCCGAGGCTGTGGGCTGTCTTAAAGTGGGAGC
 AGCTAATCAGATCGCCAGGCTTGGCCCTCAGAAAAAGTGTGAAACGTCGCTTGCACCGGACGCTCAAGCCC
 TGCAGCATCTGCTGGACTCAGGCGTGTCTTGTGCTGTTCTGGGAGGCTTGGCCACATCCCTCATGGCCCAT
 TTTATCCCCGAAATCTCGGTGTGTCAACAGTGTAGAGGCTGGGAGGAGTGTCTCACCCTGAGCTGACTTTGTT
 CTCTCTTCAAACTTCTGAGCCTTCTTGGGATTCTGGAAGGAACCTCGGCGTGAAGAACATGTGACTTCCCTT
 TCCCTTCCACTCGTGTCTTCCACAGGCTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCCTCGCTTCC
 CAAGTTAGCAGGTGTCTGTGTGTTTCAAGTGGAGGACATGTGAGTCTTGGCAGAAAGCATGGCCCATGTGCATCTCAGGA
 GGAGGACAGAAAGCCAGCTCAGTGTGAGTCTTGGCAGAAAGCATGGCCCATGTGTGCATCTCAGGAGGAG
 ACAGAAAGCCAGCTCAGTGGCCCCGCTTCCCAACCCCAAGCCGAGCAGGGCAGAGCAGCCCTCTCTG
 GAAGTGTGTCAAGTCCGATTTGAGCCTTGTCTGGGGCCAGCCCAACACCTGGCTTGGGCTCAGTGTCTGA
 GTTGCAGTAAGCTATACTTGAATCAAA

FIGURE 64

MTTWSLRRRPARTLGLLLLVLVGLFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW
IFGGSIHVFRVPREYWRDRLLKMKACGLNTLTITYVPWNLHEPERGKFDGSGNLDLEAFVIMA
AEIGLWVILRPGPYICSEMDLGGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLSRVVPLQ
YKRGPIIAVQVENEYGSYNKDPAYMPYVKKALEDGRGIVELLTSDNKDGLSKGIVQGVLAT
INLQSTHELQLLTTFLPNVQGTQPKMVMEYWTGWFDSWGPHNILDSEVLKTVSAIVDAGS
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAKYMKLRDFFGSISGIP
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGNGQSFYIILYE
TSITSSGILSGHVHDRGQVFNVTVSIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN
IDDQRKGLIGNLYLNDSPKFNRIYSLDMKKSFFQRFGLDKWXSLEPETTLPAFFLGSLSIS
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSSGINQVIVFEETMAGPA
LQFTETPHLGRNQYIK

Signal sequence:

amino acids 1-27

Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

N-myristoylation site.

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCAGGACC
CTGGTGAGGGTTCTCTACTTGGCCCTTCGGTGGGGGTCAAGACGCGAGCACCTACGCCAAAGG
GGAGCAAAGCCGGGCTCGGCCGAGGCCCCAGGACCTCCATCTCCCAATTGTGGAGGAATC
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCC**ATG**GGCTCCC
AAGAAGCTGTCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGAGCGCTACTGCTGCCCA
GGCAGACACTCGGTGCTTCGTAGTGGATAGGGGTATGACCGGTTTCCTTAGACGGGGGCC
CGTTCGCGTATGTGTTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC
CGGCTTTTGAAGATGCGATGGAGCGGCCCTCAACGCCATACAGTTTATGTGCCCTGGAACTA
CCACGAGCCACAGCCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA
ATGAGGCAGCTCTAGCGAACCTGTGGTCATACTGAGACAGGACCTTACATCTGTGCAGAG
TGGGAGATGGGGGCTCTCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC
AGATCCAGACTTCCTTGCCGAGTGGACTCCTGGTTCAGGTTCTTGCTGCCCAAGATATATC
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTCAAGTGGAGAATGAATATGGTAGC
TACAGAGCCTGTGACTTCAGCTACATGAGGCACCTGGCTGGGCTCTTCCGTGCACCTGCTAGG
AGAAAAGATCTTGCTCTTACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG
GACTCTATACCACTGTAGATTTTGGCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
CTGGGGCCAGAATCACTCCACACGGTCTGTGTGAGCTGTAAACCAAGGACTAGAGAACATGC
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTTG
AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC
TATATCTGAAGCAGGGGACCCACACCTTAAGCTTTTGTCTCTCGAGATGTCATCAGCAAGT
TCCAGGAAGTTCCCTTGGGACCTTTACCTCCCCGAGCCCCAAGATGATGCTTGGACCTGTG
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT
TCATTCAATCTTGCCAAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC
GAACCTATATGACCCCATACCAATTTTGAAGCCAAACACCTTCTGGGTGCCAAATAATGGAGTC
CATGACCGTGCCCTATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG
AGACAAACTATTTTGAAGGGGAACTGGGGTCCAAACTGGATATCTTGGTGGAGAACATGG
GGAGGCTCAGCTTTGGGTCTTAACAGCAGTGACTTCAAGGGCCGTGTTGAAGCCACCAATCTG
GGGCAAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG
GTGGTTTCCCTCCAGTTGCCAAATGGCCATATCTCAAGCTCCTTCTGGCCCCACATTTCT
ACTCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACCAATTTCTATATCTACCTGGATGG
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGATCTGGACAAAGCAGGGGCC
ACAACAGACCCCTCTACGTGCCAAGATTCTGCTGTTTCTTAGGGGAGCCCTCAACAAAATTA
CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTGGATAAGCCTATC
CTCAATAGCTAGTACTTTGCAAGGACACATATCAATTCCTTTGAGCTGATACACTGAG
TGCTCTGAACCAATGGAGTTAAGTGGGCACTGAAGGTAGGCGGGCATGGTGGCTCATGC
CTGTAATCCAGCACCTTTGGGAGGCTGAGACGGGTGATTACCTGAGGTGAGGACTTCAAGA
CCAGCTGGCCAACTGGTGAAACCCGCTCTCCACTAAAAATACAAAATTAGCCGGGCGTG
ATGGTGGGCACCTCTAATCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC
AGGAGGCAGAGGTTGCAGTGAAGTGGAGGTTTGACCACTGCACTCCAGCCTGGCTGACAGTGA
GACACTCCATCTCAAAAAAAAAA

FIGURE 66

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDGRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGPLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFS YMRHLAGLFRALLGEKILFTTDGPEGLKCGSLRGLYTTVDGFPADNMTKIF
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSSTRSVSAVTKGLENMLKLGASVNMVMFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPSPKMML
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIPEPTPFWVFN
NGVHDAYVMVDGVFQGVVERNMRDKLFLTGLGSKLDILVENMGRLSFGSNSSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWFWPLQLPKWPYPQAPSGPTTFYSKTFPILGSGVGDTFLYL
PGWTKGQVWINGFNLGRYWTQKGPQOTLYVPRFLFPRGALNKITLLEEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

Signal sequence:

amino acids 1-27

N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 4-8

Casein kinase II phosphorylation site.

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,
603-607, 644-648

Tyrosine kinase phosphorylation site.

amino acids 191-198

N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

FIGURE 67

GCTTTGAACACGTCCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC
ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTGAGGATACCTTTGAAGGAATATT
CTTTGAAAAAGTCAGAGAAGAGAGCAGTTTGTAGTACATTCCAGATGTCAAAAACGATTTT
GCGTTCTTCTTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT
CTTGTGAGAAGTTAGTAAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG
AAAAACTCAGGCAGCACATTTACGCAACGCCAGGACAAGCAGGAGTTGCATCTGTTTCATG
CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC
AATTCAGAAGCTAAATTCCTGCTAAGATTCTCAAATGACTAACCTCCAAGAGCTCCACC
TCTGCCACTGCCCTGCAAAAAGTTGAACAGAGCTGCTTTTAGCTTTCTTCGCGATCACTTGAGA
TGCTTTCACGTGAAGTTCACATGATGTGGCTGAAATTCCTGCCTGGGTGTATTGTCTCAAAAA
CCTTCGAGAGTTGTACTTAAATAGGCAATTTGAACTCTGAAAAACAATAAGATGATAGGACTTG
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAAGTT
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTACATAATGACGGCAC
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACCTCAGA
ACTGTGAGCTAGAGAGAATCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACCTGGAT
TTAAAGTCCAATAACATTGCGACAATTGAGGAAATCATCAGTTTCCAGCATTTAAACGACT
GACTTGTTTAAATTTATGGCATAACAAAATTGTTACTATTCCTCCCTCTATTACCCATGTCA
AAAACTTGGAGTCACTTTATTTCTCTAACACAAGCTCGAATCCTTACCAGTGGCAGTATTT
AGTTTACAGAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTTCAATAGA
AATAGGATTGCTTCAGAACCTGCAGCATTTCATATCACTGGGAACAAAGTGGACATTCTGC
CAAAACAAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAAGTGCATCACC
TCACTCCAGAGAAAGTTGGTCAGCTCTCCAGCTCACTCAGCTGGAGCTGAAGGGGAAGT
CTTGGACCGCTGCCAGCCAGCTGGGCCAGTGTGCGATGCTCAAGAAAAGCGGGTCTGTTG
TGGAAGATCACTTTTGTATACCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA
AATATTCCTTTGCAAAATGGGATTTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC
AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG
ATACATCTTTTAAAAATAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT
GTTCAATGTTTGTAGGGTTTAAAGTCATTCATTTCCAATCATTTTTTTTTTTCTTTGGGG
AAAGGGAAGGAAAAATATAATCACTAATCTTGGTCTTTTTAAATGTTTGTAACTTGGAT
GCTGCCGTACTGAAATGTTTACAAATGCTTGCTGCTAAAGTAAATGATTAAATTGACATT
TTCTTACTAAAAA

FIGURE 68

MAYMLKKLLISYISIIICVYGFCIYTLFWLFRIPLKEYSFEKVVREESSFSDDIPDVKNDFAPL
LHMVDQYDQLYSKRFGVFLSEVSENKLR EISLNHEWTFEKL RQHISRNAQDKQELHLFMLS
VPDAVFDLTDLVDLKLLELPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH
VKFTDVAEIPAWVYLLKNLRELYLIGNLSENKMMIGLESRLRLHLKILHVKSNTLKVPSN
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCLELIPHAIFSLSNLQELDLKS
NNIRTIEEIIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESYFSNNKLES LPPVAVFSLQ
KLRLCDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITSLP
EKVGQLSQTQLLELKGNCLELRLPAQLGQCRMLKKSGLVVEDHLFDLPLEVKEALNQDINIP
FANGI

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 241-245, 248-252, 383-387

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 326-330

Casein kinase II phosphorylation site.

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

FIGURE 69

CCCACGCGTCGGGCCCTTCTCTCGGACTTTGCAATTTCCATTCCCTTTTCATTGACAAACTGACTTTTTTTTATTCTCT
 TTTTTTCATCTCTGGGCCAGCTTGGGATCCTTAGGCGCCCTGGGAAGACATTTGTGTTTTACACACATAAGGAT
 CTGTGTTTGGGGTTTCTCTCTCCCTGCCATTGCGATTGCTTAGTGGTTGTGTGGGGAGGGAGACACGCTGG
 GCTAGCTCTTGTCTTGCACCTATGCTGCTAGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCTCTGT
 ATGCTGTGTGATCTCTGGGGCCCTTGCTCTCTGCTGATAGTTGTCTGTGCTCTGTCTTTACTTCAAAATACACAAC
 GCGCTAAAGACTGCAAGAGAACCTGAGGCTGTGGCTGTAAAAAATCACACCCAGACAGGTTGTGGTGGGGCCAG
 AACAGCCAGGCCAAACCATTTGCCACGGAGTCTTGTCTGCCCTGCAGTCTGTGAAGGATATAGAATGTGTGCC
 AGTTTGTATTCCTGTCACCTTTGCTGTTGCGACATAAATGAGGGCTCTGAGTTAGGAAAGGCTCCCTTCTCAA
 GCAGAGCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTGTGATGTGCAAGGACAGAAAGGACAG
 CTCCCATCAGTTTTCATGAAATAAATCAGTGCTGCTGGGAAACAGCTGCTGGAGATCCCTACAGAGAGCTTC
 CACTGGGGCAACCCCTTCAGGAAGGAGTTGGGGAGAGAGAACCCCTCACTGTGGGAATGCTGATAAACCACTCA
 CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGTGGAACTGACGCTTTCCCTGGAGGTGTCCAGAAA
 GCTGATGTAAACACAGAGCCTATAAAAGCTGTCGGTCCCTTAAGGCTGCCAGCGCTTGCCAAATGGAGCTTGT
 AGAAGGCTCATGCCATTGACCCCTCTTAATTTCTCTCTGTTTGGCGGAGCTGCAATGGCGGAGGCTGAAGGCAAT
 GCAAGCTGCACAGTCACTTAGGGGGTGCCATATGGCAGAGACCACAAAGCCATGATCCTGCAACTCAATCCC
 AGTGAGAACTGCACTTGACAAATAGAAAGACAGAAAAACAAAGCATCAGAAATATCTTTTCTATGTCCAGCTT
 GATCCAGATGGAAGCTGTGAAAGTAAACATTAAGTCTTTGACGGAACCTCCAGCAATGGGCTCTGTAGGG
 CAGTCTGCAGTAAAAACGACTATGTTCTGTATTGAAATCATCATCCAGTACATGACGTTTCAAATAGTTAGT
 GACTCAGCAAGAATTCAGAACTGTCTTTGTCTTCTACTACTTTCTCTCCTAACATCTCTATTCCAAACTGT
 GCGGTTTACCTGGATACCTTGGAGGATCCTTACCAGGCCCAATTACCCAAAGCCGATCCTGAGCTGGCTTAT
 TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTCAAAGAGATTTTCTCAGAAATAGAC
 AACAGTGCRAATTTGATTTTCTTGCCATCTATGATGGCCCTCCCAACTCTGGCCGTATGGCAAGTCTGT
 GGCCTGTGACTCCCACTTCGAATCGTCATCAAACTCTCTGACTGTGCTGTGTTGCTACAGATTATGCCAATCT
 TACCGGGGATTTTCTGCTTCCATACACTCAATTTATGACGAAAAATCAACACTACATCTTTAACTGTCTCTCT
 GACAGGATGAGAGTTATTATAGCAAACTCCATCAGAGGCTTTTAACTTAATGGGAATAACTGTCAACTAAAA
 GACCCACTTGCAGACCAAAATATCAAAATGTGTGGAATTTTCTGTCCCTCTTAATGGATGTGCAACTCAGA
 AAGGTAGAAGTCACTCAATATTCTTACCAAAATATAACCTTTCTGCATCCTCAACTCTGGAAGTGATCACC
 CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGACATAATTCTCAGTGGAGATAATATACATAACA
 GAAGATGATGTAAACAAAGTCAAAATGACCTGGGCAAAATAACACACGATGCGCTCTTTTGAATCCAAATCA
 TTTGAAAAGACTATCTTGAATCACCATATTATGTGGATTGGAACCAAACTCTTTTGTTCAGGTTAGTCTGCAC
 AACTCAGATCCAAATTTGGTGGTGTCTTGATACCTGTAGAGCCTCTCCCACTCTGACTTTGCATCTCCAACC
 TACGACCTTAATCAAGAGTGGATGTAGTCGAGATGAACTTGAAGGTGATCCCTTATTGGACACTATGGGAGA
 TTCCAGTTTAAATGCTTTAAATTTCTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAAGTTTGTATATGTAT
 AGCAGTGACCCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAACGAGACATTTCTCATATAAATGG
 AAAACAGATTCCATCATAGGACCAATTCGCTCGAAAAGGAGTCGAAGTGCAGGTGGCAATTCAGGATTTCAAGAT
 GAAACAGATCCGGAAGAACTCCAAACAGCCTTCAACAGTGTGCTCTGTTTTCCTTATGGTTCTAGCTCTG
 AATGTGGTGACTGAGCGACAATCAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCGAAGCTG
 CAGAATATTAACATAAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAGGAAATGCTACCTCGT
 GGCTACACATATTATGAATAAATGAGGAAGGCCCTGAAAGTGACACAGGCTGCTGTAAAAAAA

FIGURE 70

MELVRRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI
ERPENKSIRIIFSYVQLDDPGSCSEENIKVFDGTSSNGPPLLGQVCSKNDYVPVFESSSSLT
FQIVTDSARIQRTVFVFFYFSPNISIPNCGGYLDLTLEGSFTSPNYPKPHELAYCVWHIQV
EKDYKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSNSLTVVLS
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVVISKSYLEAFNSNGNNLQLKDPTCRP
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGNST
VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN
LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSVYL
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDIIIGPIRLKRDRSASGNSGFQHETHA
EETPNQPFNSVHLFSFMVLALNVTVATITVRHFVNQRADYKYQKQLQNY

Signal sequence:

amino acids 1-24

Transmembrane domain:

amino acids 571-586

N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,
394-398, 419-423

Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,
408-412, 463-467, 520-524, 556-560

Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
GGACATGCGGCCCCAGGAGCTCCCGAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGTTGCTGC
TGCTGCCGCCGCCGCGCTGCCCGCCAGCGCCAGCGCTTCGACCCCACTGGGAGTCC
CTGGACGCCCGCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG
GGGAGTGTTTCCGTGCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAAGGAAA
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCTCCTAGTTTCAAATATGAAGAT
TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAAGTGGGCAGATATTTTTCAGGC
CTCTGGTGCCAAATACATTGTCTTAACCTCCAAACATCATGAAGGCTTTACCTTGTGGGGGT
CAGAATATTCGTGGAACGTGAATGCCATAGATGAGGGGCCCAAGAGGGACATTGTCAAAGGAA
CTTGAGGTAGCCATTAGGAACAGAACTGACCTCGCTTTGGACTGTACTATTCCTTTTGA
ATGTTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT
ATATAATGAAAGCCCGATTTCGGGGCACAGTAGTACCAATGATCGTTGGGGAGCTGGTAGCA
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA
CATAAATGGGAAAACATGCATGACAATAGACAAACTGCTCCTGGGGCTATAGGAGGGGAAGCTGG
AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTTCATGTG
GAGGAAATCTTTTGATGAATATTGGGCCCACTAGATGGCACCATTCTGTAGTTTTTGAG
GAGCGACTGAGGCAAGTGGGGTCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA
TACCTGGCGATCCCAGAATGACACTGTCAACCCAGATGTGTGGTGACACATCCAAGCTTAAAG
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCACTCAGGACAGCTGTTCTTGGCCAT
CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCAATTCATCAGATGC
CGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG
ATGCTGCAAGTTATGTCTAAGGCTAGGAACATCAGGTGTCTATAATTGTAGCACATGGAGA
AAGCAATGTAACTGGATAAGAAAATTTATTTGGCAGTTCAGCCCTTTCCCTTTTCCCACTA
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCATTAAAGTC
TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTACATTTATAGTAG
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCATG
ATTATATAGTTATGCATCACTTAATATGGGGATATTTCTGGGAAATGCATTGCTAGTCAAT
TTTTTTTTGTGCCAACATCATAGAGTGTATTACAAAATCCTAGATGGCATAGCCTACTACA
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG
AATACGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG
GTACAGTAAAAACTGTGTAATAAATAAATGGTGCACCTGTATAGGCGCACTTACCACGAATGGAG
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCATAGGACATTA
TTGAACACTGCCAGACGTTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA
GTTTTCTCTTCTCAATTATAAATAACATAAGTGTACTGTAACCTTACAAACGTTTAAATT
TTTAAAACCTTTTGGCTCTTTTGTAAATAACACTTAGCTTAAAACATAAACTCATTGTGCAA
ATGTAA

FIGURE 72

MRPQELPRLAFPLLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG
VFSVPFSGSEWFWWYQKEKIPKYVEFMKDNYPSPFKYEDFGPLPTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRDRLRFGLYYSLEW
FHPFLFEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESFVRGTVVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVQQLVETVSCGNNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVPDVPWYTSKPKKEKLVAIFLKWPTSGQLFLGHKPAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKKGWALALTNVI

Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,
375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

FIGURE 73

AGCAGGGAAATCCGGATGTCCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT
 TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC
 TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTG
 CTTCTCACTTCCATCTCGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAAGACT
 TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAGGCCAACCAGC
 AGCTGAATTTACAGAAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG
 GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGA
 TGGATTCTGGTTCATCTCTAGGATTAGCCCAAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG
 TCCTGATTGTGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGAT
 ACTTGGACTAATCTGTCGATTCAGAAATTATCACCACCAAGATCCCATATTCACACTCA
 AACTGCAACACAACAACAGAATTATTGTGTCAGTGACAGTACCTACTCGGTGGCATCCCTTT
 ACTCTACAATACCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG
 AGAAAAAATTGATTGTGTACAGAAGTTTATTGGAAGTACGACCATGTCTACAGAAAC
 TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAAATGAAGCTGCTGGGTTTGGAGGTGTC
 CCACGGCTCTGCTAGTGCTGTCTCTCTCTTTGGTGTGTCAGCTGGTCTTGGAATTTTGC
 TATGTCAAAAGGTATGTGAAGGCCTTCCCTTTTACAAAACAAGATCAGCAGAAGGAAATGAT
 CGAAACCAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCCTAATGAGGAATCAAAGA
 AAAGTGATAAAACCCAGAGAGTCCAGAGTCCAAGCAAAACTACCGTGCATGCTCGGAA
 GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACTGAGGCTGGTTTCTTTCATGCTCC
 TTACCTGCCCCAGCTGGGGAATCAAAGGGCCAAAGAACCAAGAGAAAGTCCACCTT
 GGTTCCTAAGTGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGACCAAGAGAAATGC
 CCTTCTCCTTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCACGGCC
 TTTCTAGCCTGGCTATGTCTTAATAATCCCACTGGGAGAAAGGAGTTTTCGAAAGTGCAA
 GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAGGCCTCCTGGCTGTCTGAGGCTAGG
 TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGAGCTCAGAC
 CCTTCTCCTCAGCTCTGAAAGAGAAACACGTATCCCACTGACATGTCTCTGAGCCCGTA
 AGAGCAAAAGAAATGGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAG
 ACCTAATCTCTGTAAGCTAAAATAAAGAAATAGAACAAAGGCTGAGGATACGACAGTACACT
 GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTCTTCTGTAACACATTGAGTTGGA
 ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT
 AGGAAATATACTTTTACAAGTAACAAAAATAAACTCTTATAAATTCTATTTTATCTGA
 GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA
 ACAACATTGCTGAATAGCTACTATATGTCAAGTGTGTGCAAGGTATTACACTCTGTAAT
 TGAATATATTCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTCTTCA
 TTTTGGATATTTCTAGCTTATCTACTTCTTCCAACTAATTTTATTTTGTCTGAGACTAATCTT
 ATTCAATTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG
 TACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCCATTAAACAAATGTATCACTA
 GCCCTCCTTTTCCAACAAGAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA
 AGCATTAGAAAACTT

FIGURE 74

MARCFSLVLLLSIWTTTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTAEAEACR
LLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGLIWKVPVSRQF
AAYCYNSSDTWTNSCIPEIIITKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAPKNEAAGFGGVPTALLVLALLFF
GAAAGLGFCYVKRYVKAPFPTNKNQKEMIETKVVKEEKANDSNPNEESKTKDNPEESKSP
SKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 75

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTGCGGTGCGCGACTTTTCACGATGG
CTGCGCCCAACCTTACTACCTTCTGTGCGCCCTGCTCTCTGCTGCGCTTCTTACTCGTGAGGAA
ACTGCGCGCGCTCTGCCACGGTCTGCCACCCAACGCGAAGACGGTAACCGTGTGACTTTG
ACTGGAGAGAAGTGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA
TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGCCCAACACAAT
TCTTTTCTTCCGCTTGGATATTGCGATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC
TGATGACGTGCAAAACCCCCCTATATATGGGCCTGAGTATATCAAGTACTTCAATGATAAA
ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA
TTGGTCTAATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAAC
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAGGTGGCAA
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG
AGGAGAATGTGATCCGAGAATTAACTTAAATGAGCTATACCAGCGGGCCAAGAACTATCA
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCAACACAGTGTGAGA
TGGGAAAAACAAGAAGGATAAATAAGATCCTCACTTTGGCAGTGCTTCCCTCTCCTGTCAATT
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTNNATNNATGTTTCCCTTTGG
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTGAG
GCACCTTACAGGAAGGCCTGCCATGCTGTGGCCAACGTTTTCACTGGAGCAAGAAAGAGATC
TCATAGGACGAGGGGGAAATGGTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC
AGCTATTGAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAG
TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCTAAGGTAAGATGCTGGGGTATAGAA
CGCTAAGAAATTTCCCCCAAGGACTCTTGCTTCTTAAAGCCCTTCTGGCTTCGTTTATGGTC
TTCATTAAAGTATAAGCCTAACTTTGTGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG
TTTTTATCATTTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAA
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCTGGACTTTCAC
TAACCCCTTGACATACTCCCCACACCCAGTTGATGGCTTTCGTAATAAAAAGATTGGGATT
TCCTTTTG

FIGURE 76

MAVLAPLIALVYSVPRLSRWLAQPYLLSALLSAAFLLVRLPPLCHGLPTQREDGNPCDFD
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAP IYADLSLKYN C
TGLNFGKVDVG RYTDVSTRYKVSTSP LTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

Signal sequence:

amino acids 1-48

Transmembrane domain:

amino acids 111-125

N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 154-158, 265-269

Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.

amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.

amino acids 244-253

FIGURE 78

MGLLLLVLPLLLPGSYGLPFYNGFYYSNSANDQNLGNHGKGKDLLNGVKLVVETPEETLFTYQ
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLAIGLRHRSFGDYQGRVHLRQD
KEHDVSLLEIQDLRLLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ
VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR
HRRLLHRYDVFCFATALKGRVYYLEHPEKLTLTAREACQEDDATIAKVGQLFAAWKFHGLDR
CDAGWLADGSVRYPPVPHPNCGPPEPGVRSFGFPDPQSRLYGVICYRQH

Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site.

amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

FIGURE 79

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC
TTCTGTGTGCTACTGAGGCACGGGGCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC
CAGGGGAGGGTGCACCAGGCGGGCCCCCTGAGCGACGCTCCCCATGATGACGCCACGGGAA
CTTCCAGTACGACCATGAGGCTTCTCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCA
CCCCAGAGGAAAGCCAGGCCCTCTGSGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC
GGCGACGGCTGGGTGTCTGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA
CATACGGGACTCGGTGAGCGCGGCCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG
GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGTGAAGAATTTTCATGAC
GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGC
CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCTCTGCACCCCGAGG
AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCCTGGAGGACCTGGACAGAAACAAA
GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA
GGAGCCGGCGTGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG
ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCC
CTGGTGAAGCCAAACCCTGTGTCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA
AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG
ACCTGACCCGGCACCACGATGAGCTGTGAGCACCCGCGCACCTGCCACAGCCTCAGAGGCCCG
CACAATGACCGGAGGAGGGGCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCGCAGGAG
GCAGATGCAGTCCCAGGCATCTCTGCCCCCTGGGCTCTCAGGGACCCCTGGGTGCGCTTC
TGTCCCTGTACACCCCCAACCCAGGGAGGGGCTGTCTAGTCCCAGAGGATAAGCAATAC
CTATTTCTGACTGAGTCTCCAGCCCAGACCCAGGGACCTTGGCCCCAAGCTCAGCTCTAA
GAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCTT
GGCCCCAGCCCTCTCTGCTGGCTGGCTGGGACACCTCTCTGTCAGGAGGCAATAA
AAGCCAGCGCCGGGACCTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 80

MMWRPSVLLLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGR
EVAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWSLAE LR AWIAHTQQRHIRDSVSAAWDT
YDTRDRGRVGWHEELRNATYGHYAPGEEFHDVEDAET YKKMLARDERRFRVADQDGDSMATRE
ELTAFLHPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAWVQTERQQ
FRDFRDLNKGHLGDGSEVGHVWLPPAQDQPLVEANHLLHESDTRDKDGRLSKAEILGNWNMFV
GSQATNYGEDLTRHHDEL

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,
291-295, 298-302

N-myristoylation site.

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

FIGURE 81

GGGGCTTGCCCTTCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG
 GCGGCGGGCGCGGGTGCAGAGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTGCTCCAG
 CCTGTCTGTGCTCGTTTGGCGCCCCGCTCCCGCGGTGCGGGGTGCACACCGATCTGT
 GGCTTCGCTCGATTGCGCGCCGAGGCGCTCCAGACCTAGAGGGGCGCTGGCTGAGCAG
 CGGGTCGTCTGTCTCTCTCTCTGCGCCGCGCCCGGGGATCCGAAGGGTGCAGGGCTCT
 GAGGAGGTGACGCGCGGGGCTCCCGCACCTTGGCCTTGCCCGCATCTCCCTCTCTCCAG
 GTGTGAGCAGCCTATCAGTCACCATGTCCGCGAGCTGGATCCCGGTCTCGGCTCGGTGTG
 TGTCTGTCTGTGCTCGCGGGGCGCGGGGCGAGGGAGCCGCTCCCATTGCTATCACATG
 TTTTACCAGAGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCAGGGGGCTGCC
 CTCTTGAGGAATTTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGG
 GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC
 TGGTCGAGAAAACTATTCTCAGTAGATGCCAATGGCATCCAGTCTCAAAATGCTTCTAGAT
 GGTCTGCTTCTTTCACAGTAACATAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA
 GTGTCCACAGCACATCCACCAAGTAAACGACTAAAGAAAAACCCCGAGAAGAAAACTGG
 CAATAAAGATTGTAAGCAGACATTGCATTCTCTGATTGATGGAAGCTTTAATATTGGGCAGC
 GCCGATTTAATTACAGAAGAAATTTGTTGGAAAAGTGGCTCTAATGTTGGGAATTGGAAACA
 GAAGGACCACATGTGGGCTTGTTCAGGCCAGTGAAACATCCCAAAATAGAATTTTACTTGAA
 AAATCTTACATCAGCCAAAGATGTTTGTGTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA
 ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATCTTCACGGTAGATGCTGGA
 GTAAGAAAAGGGATCCCAAAGTGGTGGTGGTATTTATTGATGGTGGCTCTTGATGACAT
 CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTATAGTTTCTGTGGCCA
 AGCCTATCCCTGAAGAACTGGGGATGGTTCAGGATGTCACATTTGTTGACAAGGCTGTCTGT
 CGGAATAATGGCTTCTCTCTTACCACATGCCCACTGGTTTGGCACCACAAAAACGTAAA
 GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT
 CAGTGAACATTGGCTTTCTAATTGATGGCTCCAGCAGTGTTGGAGATAGCAATTTCCGCTTC
 ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAATCTCGGACATTGGTGCCAAGAT
 AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACCA
 AAGAGAATGTCCTAGCTGTCTATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT
 GATGCCATTTCTTCTCACTGTGTAGAAATGTGTTTGGCCCTATAAGGGAGAGCCCCAACAGAA
 CTTCTAGTAATTGTGCACAGATGGGCAGTCCCTATGATGATGTCCAAGGCCCTGCAGCTGCTG
 CACATGATGCAGGAATCAGTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG
 AAAGATATGGCTTCTAAACCGAAGAGTCTCACGCTTCTTTCACAAGAGAGTTACAGGATT
 AGAACCAATTGTTTCTGATGTCTCAGAGGCATTGTAGAGATTCTTGTAGAAATCCAGCAAT
 AATGGTAACATTTGACAACTGAAGAAAAAGTACAGGGGATCCAGTGTGTAAATTGTATCT
 CTCTAAATAGTGAATGCTTTAGCATACTAGAATCAGATACAAAATATTAAAGTATGTCAAC
 AGCCATTTAGGCAAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT
 ACTTTGTTAAAAACACTGCTGAGGCTCTAATCATGGCTCTTAGAACTCAGGAAAAGAGGA
 GATAATGTGGATTAAAAACCTTAAGAGTTCTAACCTGCTTAACCTAATGATACAGATATGCAAA
 TTCCATAGCTCAATAAAAGAATCTGATACTTAGACCAAAAAAAAAA

FIGURE 82

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPPGGCPLEEFVSY
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLNFTSAKDVLFAlKEVGFRRGNSNTGKAL
KHTAQKFPTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSAKPIPEELG
MVQDVTFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI
DGSSSVGDSNFRMLLEFVSNIAKTFEISDIGAKIAAQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRES PNKNFLVIVTDGQSYDDVQGPAAAHADAGITI
FSVGVAWAPLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

Signal sequence:

amino acids 1-24

N-glycosylation site.

amino acids 100-104, 221-225

Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,
425-429, 478-482, 528-532

N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.

amino acids 145-149

FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCGAGCCC
GGCGGCGCTCCCGCGGGGAGCGAGCAGATCCAGTCCGCGCCCGCAGCGCAACTCGGTCCAGTCG
GGGCGGCGGCTGCGGGCGCAGAGCGGAGATGCGAGCGGCTTGGGGCCACCTGCTGTGCTGCTGC
TGCTGGCGGCGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTC
AAGCCCGGCGCGGCTCTCAGCTACCCGAGGAGGAGGCCACCTCAATGAGATGTTCCGCGA
GGTTGAGGAACCTGATGGAGGACACGCGACACAAATTGCGCAGCGCGGTGGAAGAGATGGAGG
CAGAAGAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCAGCTAT
CACAAATGAGACCAACACAGACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAAT
TCACAAGATAACCAACAACAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCATCATCTG
TGGGAGACGAAGAAGGCAGAGGAGGCCACGAGTGCATCATCGACGAGGACTGTGGGCCAGC
ATGTACTGCCAGTTTTCGCGACTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCT
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA
TGGCCACAGGGGCGAGCAATGGGACCATCTGTGACAAACAGAGGGACTGCCAGCCGGGGCTG
TGCTGTGCTTCCAGAGAGGCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACTGGGAGCTAGAGCCTGATG
GAGCCTTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG
GTGATGTGTGTCGAAGCCGACCTTCGTGGGAGCCGTGACCAAGATGGGAGATGCTGTGCTGCC
CAGAGAGGTCCTCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG
AGGACCTGGGAGGAGGCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCT
GCACTGCTGGGAGGGGAAGAGATTAGACTCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA
ATAGCTAATTTATTTCCCGAGGTGTGTGCTTTAGGCGTGGGCTGACAGGCTTCTTCTCTACA
TCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCAATTTGTTCCG
TCCCCAGGCTGTTCTCCAGCTTTCACAGTCTGGTGTCTGGGAGAGTCAGGCAGGGTTAAAC
TGCAGGAGCAGTTTGCACCCCTGTCCAGATTATTGGTGTCTTTGCTCTTACAGTTTGGCAG
ACAGCCGTTTGTCTACATGGCTTTGATAATTGTTGAGGGGAGGAGATGGAACCAATGTGG
AGTCTCCTCTGATTGGTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAACATCAA
CCTGGCAAAATGCAACAAATGAATTTCCACGCGACTTCTTCCATGGGCATAGGTAAGCTG
TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTCACCCTGCATTACATGTTATTCATCC
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC
CCTCTCTCAGCACAGCTGGGGAGGGGGTTCATGTTCTCCTCGTCCATCAGGATCTCAGAG
GCTCAGAGACTGCAAGCTGTCTGCCAAGTCAACAGCTAGTGAAGACCAGAGCAGTTTCAT
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCACACCCAGCCTTGGTGCCACCAA
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTCTTGAGGCATGCACATGTGGAATTAAG
GTCAAACTAATTCTCAGATCCTCTAAAGTAACTACTGTTAGGAACAGCAGTGTTCTCAC
AGTGTGGGGCAGCCGCTTCTTAATGAAGACAATGATATGACACTGTCCCTCTTGTTCAGT
TGCATTAGTAACCTTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA
GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC
TGAAAGACAATTATCAACCACGCTGGAGAAAAATCAAACCGAGCAGGGCTGTGTGAACATCGGT
TGAATATGCGACTGCGAACAACCTGAACCTACGCCACTCCACAATGATGTTTTCAGGTGTCA
TGGACTGTTGCCACCATGTATTATCCAGAGTCTTAAAGTTTAAAGTTGCACATGATTGTA
TAGCATGCTTTCTTGTAGTTTAAATATGTATAAACATAAGTTGCATTTAGAAATCAAGC
ATAAATCACTTCAACTGCAAAAAA

FIGURE 84

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYQEEATLNEMFREVEELMEDTQ
HKLRSAVEEMEAEEAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQFCRGQRMCLTRDSECCG
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCACAFQRLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEET

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

Order	Product	Price	Quantity	Total
1	1000	1000	1	1000
2	2000	2000	1	2000
3	3000	3000	1	3000
4	4000	4000	1	4000
5	5000	5000	1	5000
6	6000	6000	1	6000
7	7000	7000	1	7000
8	8000	8000	1	8000
9	9000	9000	1	9000
10	10000	10000	1	10000

21-21-11

FIGURE 86

MRLLVAPLLLA WAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVCNDLFLTA
VPPALPAGTQTLLQLQSNISIVRVDQSELGYLANLTELDLSQNSFSDARDCDFHALPQLLSLHL
EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE
MLPNLEILMIGGNKVDAILD MNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ
LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELVSIKDFALVNLP
ELTKLDITNNPRLSFIHPRAFHHLPQMETMLMNNALSALHQQTVESLPNLQEVGLHGNPIR
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPFREMTHCLPLISPRSFPPSLQ
VASGESMVLHCRALAEPEPEIYWVTPAGLRRLTPAHAGRRYRVYPECTLELRRTAEEAGLYT
CVAQNLVGADTKTVSVVVGRALLQPRDEGQGLELRVQETHPHYHILLSWVTPPNTVSTNLTW
SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS
CHRALGDRPGLIAILALAVLLAAGLAHLGTGQPRKGVGGRRPLPPAWAFWGWSAPSVRVV
SAPLVLPWNPGRKLP RSSEGETLLPPLSQNS

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 629-648

N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 485-489

Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146,
243-247, 313-317, 488-492, 700-704

Tyrosine kinase phosphorylation site.

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

[illegible]

CTAACGACAGGGCGCTGTCTGAGAAAGGTGAAGAAGTTCGCGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT
CTAATCGCGGACGACCATCATTAAGGTGATCAAGTTCATCTCATCATCTGTCAACCGCTGTCTATACGCTGACACAA
CATCAAGTTCGACGTGGACTGCAACGCTGGACATTGAGAGGCTGACGCGGTACCGCATACCGCTGTGGCCACC
CTCTGGCACATCTTCAAGATCTCTGGGCTCTCTTACATCACTGAGTGTATGTCATCTACGCGCTCATCTGTCATGTGTA
CACACTGTGGTGGATGCTACGCGGCTCCCTCAAGAGTACTCGTTTGAGTGCATCTGTCGAGGAGAGCAGCTACGAC
CGACATCCCCGAGCTCAAGAACGACTCGGCTCTCATGCTGACCTCAITGACAATCAACGACCGCTCTATCTCCAA
GGCTTCGCGGCTCTCTCTGTGGAGGTGAGTGAGAACAGCTGCGGCGACTGAACCTCAACAAACAGGTGAGCGCT
GGACAGCTCTCGGCGAGCGGCTCTCAACGAAAGCGCGAGAACAGCTGAGACTGCACTGTCATTGTTCATGCTCAGTGGGCA
CCCTGACAGTGTGTGATCGCTGTGGAGCTGGAGGCTCTCAGCTGAGAGCTGACATCCCGGACGTGACCATCTCGCGC
CAGCATTCGCGACCTCAACGCGCTCAAGGAGCTGTGGCTCATACACACAGCGGCGAAGTTGAAGCGCTCGGCT
GGCCTCTCTGTGCGAGAACTCGCGGCGCTGCACATCAAGTTACACGACATCAAGGATTCGCGTGTGGATGTATG
TAGCCTGAAGACATGCGAGGAGCTGCACCTGACGGGCAACTGAGCGGGAGAACACACGCTCATATGCTCATGCA
CGGCTGTGGGAGCTCAAAAGCCTCAAGGTCGCTGGGCTCAAGAGAACTAAGACGACTGCCACAGGTGTGTCA
AGATGTGGCGCTGACATCGAAGAGCTGTCATCAACATGAGGGGCAACGAGCTCATGCTCTCAACAGCGCTCAA
GAAGATGGGGAACCTGACTGAGCTGGAGTGTGCTGCGTGCAGACTGGAGCGCATCCCCACCTCATCTTCAGCT
CCACAACTCTCAGAGAGATGACTCAAGGACAAACAACTCAAGAACATCGAGGAGATCATGCTTCCAGCACTCT
GCACGGCTTCACTGCTTAACTGTGGTACCAACATCTGCGCTCATATCCCATTCAGATTCGCGAACCTCATCAAA
CCTGGGCGGCTCTACTCGAACCGCAACAGATCGAAGAGTCCCCACCGACTTCTACTGCGCAAGCTCGCT
CTAGCTGGACCTGAGCACAAACAACTGACCTTCTCTCGCGACATCGGCTCTGCGAAGCTCCAGAACTCCAGAACT
AGCCATCAGCGGCCACCGATCGAAGCGCTCCTCCGAGACTCTTCTCAGTGC CGGAGCTGCGGCTCTGCACTCT
GGGCACAACTGTCGTGACTCACTGCCCTCAGGGTGGGAGAGTGACAACTGACGACGATGAGCTCGCGGG
GACCTGGTGTGAGTGCCTGCCCTGAGACTGGGCGAGTGCCCATGCTCAAGCGCAGCGGCTGTGGTGTGGAGGG
CGACCTGTTCAACACTCGCTCCGACCTGGAGGCTGAAGGAGGCGGCTGTGGAGGCTGACAGGAGCAGCGCTCGAGCGAG
GCGCGCCGACGACGAGGACGCTCGCGGCTGCGGAGCTCTGCGGCGCGAGGAGGAGGCTAGCTTCTCCGAC
AACTCCGAGCAGCGACGAGACGCTCGCGGCTGGCGAGGAGCTGTTGAGTCAGGCGACGAGCTGAGCTGCGGAG
GACAGATCTGATGCTGGGCTGGCGCTTTTCTCCCTGAGACTCAGCTCCCCAGGCAAGTGTCTGTGGAGAG
AGCAAGCTCTCAAGAGCGAGTATTTGGAATCATCAGGCTCTCTCTCTGAGGCGACCTCTGCGCCGAGCGGTAG
CTGCCACGAGAGGCTCTGGGAGCCTCACTTAGTCTCTGTTATTTATTTTCTCCATCTCCCACTGCTCTCATC
AGATAATCTATACATCTCCGAAGAATTCAGCCGAGATGGAAGGTGTTCAGGGAAGGTGGGCTGCTTTCTCCCT
TGTGCTTATTATAGCATGCGCGGGGCTATTAAACAGCACTGGATCTCATGACAGATGTGTCGGGCGAACACG
CAATGGGACGCTGACCCAGACGCTGCGCGGCTGGGCTCTGCGGTTGAGTTCTCAGCGGAGCAGCGCTCCAGCTGG
GACCGAGCGCTGGAGCTTGCCTCTCAGTTTGTGAGGATTGTGTTTGTGTTTGTGTTTTTTTTTAAACAA
AAACAACTTTTAAAGAAAGCTTTGAAATGATAGTTTGGGATATAAAGAAAGAAAGAACTTAAAGAAA
AAAGACACTATTTTCCCAAGATGATGTGAGATCTCAGGCGAGGTTGGGAGCTTCTCTGAGCAACAGCAGACGCT
TGAATCTGTTTCTCTTCTCTCGGCGCAGGTTGCGAGGTTGCTTCGAGATCTGTGTGACCTTGTGTCAGGATTC
CTATTGTTCTGGGAGGAGGTTTTTGTGTTGTTTGTGTTTTTGTGTTTCTGTTTCTTCTCTCTCTCT
ATGTGTTCTTGCGGACCATCTATTCTGTGGCTGTGCGGCGAGGAGGATTTCTGAGCTGTCCAAGGAGGAGGAG
ACTCGGTTGCTGATATCCCGGATGAAGCGTGCTCATTCGCACTCCCTCTCTGCTGCTCCGCTGCTCTCCCA
CGCAGATGTTTAAGGAGACGAGAGGAGCTCATCGCCGACATCTTGTGTTCCCACTCTCTGCGGCTGGGTTGT
CCAGTGGCACGCTGGCCTCGCTGCTTCCATCAGCGCTGTGCGCACCTGTGCTCTCAAGAGCAGACATCTTA
GAGGCTGTGTCGGGAATGGGAGGTTGCCCTCGGGAGGCGGCGCTGTGTTCCAAGCCGCTTCCGCTCTCGGCG
CTGGAGTGCACAGGAGGACTCGGCACCTGTGGTGTGGAAGCAACTGCTCTTAGATCATCTGGGTCCCACTCT
AGAAGGCTCCCGGCTTAGATCAATCAGTGGACATGAAGCACTGTTTGAAGTCTCTTGTCTTAATGATTATGT
CCATCGCTGTGCTGCTCAATTGTGTTTCTGCGCTGTGCTCATTTGATATAATCTCAGAAATATGACATGAC
CTCTCGACACCACTGAAGCAAAATCGGTTACATGTGGGTCTGAACCTGTGAGACTCGGTACATGATCAAAATAA
ATCTTAAACAGAAAATAAAAAA

FIGURE 88

MRQTI IKVIK FII IICYTVYYVHN IKFDV DCTVD IESLTGYRTYRCAHPLATLFKILASFYI
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK
RFAVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLELHLFMLS GIPD TVFDLVELEV
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAF LRENLRALHIKFTDIKEIPLWI
YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKL PQVVTDVGVHLQKLSI
NNEGTKLIVLNSLKKMANL TELELIRCDLERIPHSIFS LHNLQETDLKDN NLKTIEBIISFQ
HLHRLTCLKLWYNHIAIYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKRLVLDL SHNLT
FLPADIGLLQNLQNLAITANRIETLPPEL FQCRKL RALHLGN NVLQSLPSRVGELTNLTQIE
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:

amino acids 51-75 (type II)

N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 85-89

Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,
398-402, 493-497

N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCA**ATG**GAGCTGGCACTGCGGCGCTCTCCCGT
 CCCGCGGTGGTTGCTGCTGCTGCCGTGCTGCTGGGCCTGAACGCAGGAGCTGTCA**TTG**ACT
 GGCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG
 TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCTGGTCAT
 GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATT**TGG**AACTTTGAGGAAATGGGC
 CCGTTGACAGTGATCTCAAACCACGAAAACCCTGGCTCCAGGCTGCCAGTCTCCTATTT
 GTGGATAATCCCGTGGGCAC**TGG**GTTCAGTTATGTGAATGGTAGTGGTGCTATGCCAAGGA
 CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCA**TTG**CCACAAG
 AATTCAGACAGTTCATTCTACATTTTCTCAGAGTCTTATGGAGGAAAAATGGCAGCTGGC
 ATTGGTCTAGAGCTTTATAAGGCCATT**CAG**CGAGGGACCATCAAGTGCAACTTTGCGGGGGT
 TGCTTGGGTGATTCTTGGATCTCCCTGTTGATT**CGG**TGCTCTCCTGGGGACCTTACCTGT
 ACAGCATGTCTCTTCTGAAGACAAAGGCTTGGCAGAGGTGTCTAAGTTGCAGAGCAAGTA
 CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT
 GATCATTGAACAGAACACAGATGGGGTGA**ACTT**CTATAACATCTTA**ACT**AAAAGCACTCCCA
 CGTCTACAATGGAGTCGAGTCTAGAA**TT**CACACAGAGCCACCTAGTTTGTCTTTGT**CAG**CGC
 CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCC**CC**ATCAGAAAGAA
 GCTCAA**AA**TATTCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA
 TGGAGGAGGACTTCATGAAGCCAGTCA**TTAG**CATTGTGGACGAGTTGCTGGAGGCAGGGATC
 AACGTGACGGTGATAATGGACAGCTGGATCTCATCGTAGATACCATGGGT**CAG**GAGGCCTG
 GGTGCGGAACTGAAGTGCCAGA**ACTGC**CTAAATTCAGTCAGCTGAAGTGAAGGCCTGT
 ACAGTGACCCTAAATCTTTGAAACATCTGCTTTTGTCAAGTCTACAAGA**AC**CTTGCTTTC
 TACTGGATTCTGAAAGCTGGTCA**TAT**GGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT
 GATGAGACTGGTGACTCAGCAAGAA**TAG**GTATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT
 TGGGGCAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT
 AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTG**CAG**AGGATAAAATCATTGTCTCT
 GGAGGCAATTTGGAAATTA**TT**CTGCTTCTTAAAAAACCTAAGATTTTTTAAAAAATTGAT
 TTGTTTTGATCAAAATAAGGATGATAATAGATATTAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 90

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPNVGTGFSY
VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSES YGGKMAAGIGLELYKAIQR
GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNGLYRE
ATELWGKAEMIEQNTDGVNFYNILTKSTPTSTMESSELEFTQSHLVCLCQRHVRHLQRDALS
QLMNGPIRKKLKIIPEDQSWGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL
IVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKS YKNLAFYWILKAGHMVP
SDQGDMALKMMRLVTQQE

Signal sequence:

amino acids 1-25

N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 101-105

Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,
187-193, 195-201, 331-337, 332-338, 360-366

FIGURE 91

GGCCGCGGGAGAGGAGGCCATGGGCGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTC
GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGCCGTTATCAGGACCATGCGGCCGA
CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCGGAACCGGGCGTTGGCCGTGGCA
GGGGAGCCTGCGCCTGTGGGATTCACACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG
CACTCACGGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG
GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCCTACTACACCCG
TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATCACCCCTATGACATTG
CCTTGGTGAAAGCTGTCTGCACCTGTACCTACACTAAACACATCCAGCCCATCTGTCTCCAG
GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA
AGAGGATGAGGCACTGCCATCTCCCCACACCCCTCCAGGAAGTTCAGGTGCGCATCATAAACA
ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG
GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCCTT
GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG
GTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACCTTTGAGTGGATCCAGAAG
CTGATGGCCCAAGAGTGGCATGTCCCAGCCAGACCCCTCTGGCCACTACTCTTTTCCCTCT
TCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTTGAGCCTACCTGAGCCCATGCAGCCTGGGGC
CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTGGTAATAAACACATTCCAGTTGA
TGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 92

MGARGALLLALLARAGLRKPESQEAAPLSGPCRRVITSRIVGGEDAELGRWFPWQGSRLRW
DSHVCVSVLLSHRWALTAAHCFETYSDLSDPGWMVQFGQLTSMPSFWSLQAYYTRYFVSN
YLSPRYLGNSPYDIALVKLSAPVTTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDI FGDMVCAGNAQGGKDACFGDSGGFLACNKNG
LWYQIGVVSWGVGCGRPNRPGVYTNISHHFEWIKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 167-171, 200-204, 273-277

Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,
259-265, 269-275

Amidation site.

amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 252-263,

Serine proteases, trypsin family, histidine active site.

amino acids 78-84

FIGURE 93

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAATGGGACTCCAAGCCTGCCTCCTAGGGCT
 CTTTGCCCTCATCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACAGCGGAGGACGC
 TGCCCCCAGGCTGGGTGTCCCTGGGCGGTGCGGACCTTGAGGAAGAGCTGAGTCTCACCTTT
 GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGACGGCTGTGTGGATCCCAG
 CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAAATGTGGCTGATCTGGTGAGGCCATCCC
 CACTGACCCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCCAAGTGCATTCT
 GTGATCACACGAGCACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC
 TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAAGTCCCCAC
 ATCCCTACCAGCTTCCACAGGCCCTTGGCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT
 TTTCCCCAACATCATCCCTGAGGCAACGCTCCTGAGCCCGCAGGTGACAGGGACTGTAGGCCCT
 GCATCTGGGGGTAACCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAGACGTGG
 GCTCTGGCCACCAGCAATAACAGCCAAGCCTGTGCCAGTTCTCTGGAGCAGTATTTCCATGAC
 TCAGACCTGGCTCAGTTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC
 CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCGGATTGAGGCCAGTCTAGATGTGCAGT
 ACCTGATGAGTGCTGGTGCCAAACATCTCCACCTGGGCTACAGTAGCCCTGGCCGGCATGAG
 GGACAGGAGCCCTTCTGTCAGTGCTCATGTGCTCAGTAATGAGTCAGCCCTGCCACATGT
 GCATATCTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA
 ACATGAGGCTCATGAAGGCTGCCGCTCGGGGTCTCACCTGCTCTTCGCCCTCAGGTGACAGT
 GGGGCCGGGTGTTGGTGTGTCTCTGGAAGACACCAGTTCGCCCTACCTTCCCTGCCTCCAG
 CCCCTATGTCAACACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCAAAATGAAA
 TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCACGGCCTTCATACCAGGAG
 GAAGCTGTAAACGAAGTTCTTGAGCTCTAGCCCCCACCCTGCCACCATCCAGTTACTTCAATGC
 CAGTGGCCGTGCCTACCCAGATGTGGCTGCACCTTTCTGATGGCTACTGGGTGGTCAGCAACA
 CAGTGCCCATTCATGGGTGTCCGGAACCTCGGCCCTTACTCCAGTGTTTGGGGGGATCCTA
 TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCTCTTGGCTTTCTCAACCCAAAG
 GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTG
 TGGATGAAGAGGTAGAGGGCCAGGTTTCTGTGCTCTGGTCTCTGGCTGGGATCCTGTAACAGGC
 TGGGGAACACCAACTTCCAGCTTTGCTTGAAGACTCTACTCAACCCCTGACCCCTTCTCATC
 AGGAGAGATGGCTTGTCCCTGCCCCTGAAGCTGGCAGTTTCACTCCCTTATCTGCCCCTGTG
 GAAGCCCTGTGAACCCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCCTGAAA
 TGCTGTGAGCTTGACTTGACTCCCAACCTACCATGCTCCATCATACTCAGGTCTCCCTACT
 CCTGCCCTAGATTCTCTAATAAGATGCTGTAAGTAGCATTTTTTGAATGCCTCTCCCTCCGC
 ATCTCATCTTTCTCTTTCAATCAGGCTTTTCCAAAGGGTGTATACAGACTCTGTGCACTA
 TTTCACTTGATATTCATCCCCAATTCACTGCAAGGAGACCTCTACTGTACCGTTTACTCT
 TTCTACCTGCATCCAGAAAACAATGGCCTCCAGTGCACTACTTCTCAATCTTGCTTTATG
 GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACCTCTGTG
 ACTACTCTTGCTCTCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGCT
 TCCATTTGTAGATTTTGTCTCTCTCAGTTTACTCATTGTCCCTGGAAACAATCACTGACA
 TCTACAACCATACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA
 TGTAAAAA

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 94

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLLENVADLVRPSPLTLHTVQKWL LAAGA QKCHSVITQDFLTCWL
SIRQAELLLPGA EFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGG LHRFPPTSSLRQR
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVSGGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDDESLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPSSSYFNASGRAYPDVAALSDGYVWVSNRPVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHAGLFDVTRGCHESCLDEEVEGQGFCSPGWDPVTGWGTPTS QLC

Signal sequence:

amino acids 1-16

N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,
521-527, 533-539, 549-555

FIGURE 95

GCCGCGCGCTCTCTCCCGGCGCCACACCTGTCTGAGCGGCGCAGCGAGCCGCGGCCCGGGC
GGGCTGCTCGGCGCGGAACAGTGCTCGGCATGGCAGGGATTCCAGGGCTCCTCTTCTCTCTC
TTCTTTCTGCTCTGTGCTGTTGGGCAAGTGAGCCCTTACAGTGCCCTCGGAAACCCACTTG
GCCTGCATACCGCCTCCCTGTCGTCTTGCCCCAGTCTACCTCAATTTAGC CAAGCCAGACT
TTGGAGCCGAAGCCAAATTAGAAGTATCTTCTTCATGTGGACCCAGTGT CATAAGGGAAC
CCACTGCCCACTTACGAAGAGGCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGG
CAGCCGCACAGAGACGCAGGTGGGCATCTACATCCTCAGCAGTAGTGGAGATGGGGCCCAAC
ACCGAGACTCAGGCTCTCAGGAAAGTCTCGAAGGAAGCGGCAGATTATGGCTATGACAGC
AGGTT CAGCATTTTGGGAAGGACTTCTGTCTCACTACCTTTCTCAACATCAGTGAAGTT
ATCCACGGGTGCACCGGCACCCTGGTGGCAGAGAAGCATGTCTCAGAGCTGCCCACTGCA
TACAGATGGAAAAACCTATGTGAAAGGAACCCAGAAGCTTCGAGTGGGCTTCCTAAAGCCC
AAGTTTAAAGATGGTGGTCGAGGGGCCAACGACTCCACTTCAGCCATGCCCGAGCAGATGAA
ATTT CAGTGGATCCGGGTGAAACGCACCCATGTGCCCAAGGGTTGGATCAAGGGCAATGCCA
ATGACATCGGCATGGATTATGATTATGCCCTCCTGGAAC TCAAAAAGCCCCACAAGAGAAAA
TTTATGAAGATTGGGGTGAGCCCTCCTGCTAAGCAGCTGCCAGGGGGCAGAATCACTTCTC
TGGTTATGACAA TGACCGACAGGCAATTGGTGTATCGCTTCTGTGACGTCAAAGACGAGA
CCTATGACTTGCCTAC CAGCAATGCGATGCCAGCCAGGGGCCAGCGGTCTGGGGTCTAT
GTGAGGATGTGGAAGAGACAGCAGCAGAGAAGTGGGAGCGAAAAATTTATGGCATTTTTTCAGG
GCAC CAGTGGGTGGACATGAATGGTTCCCACAGGATTTCAACGTGGCTGT CAGAATCACTC
CTCTCAAATATGCCCAGATTTGCTATTGGATTAAAGGAACTACCTGGATTGTAGGGAGGGG
TGACACAGTGTTCCCTCCTGGCAGCAATTAAGGGTCTTCATGTCTCTATTTTAGGAGAGGCC
AAATTGTTTTTTGT CATTGGCGTGACACGTGTGTGTGTGTGTGTGTGTGTGAAGTGT
CTTATAATCTTTTACCTATTTCTTACAATTGCAAGATGACTGGCTTTACTATTTGAAAATG
GTTTGTGTATCATATCATATATCATTTAAGCAGTTTGAAGGCATACTTTG CATAGAAAATAA
AAAAA TACTGATTTGGGGCAATGAGGAATATTTGACAATTAAGTTAATCTTCACGTTTTTG
CAAAC TTAGATTTTATTT CATCTGAAC TTGTTTCAAAGATTATATTAATATTTGGCATA
CAAGAGATATGAAAAAAAAAAAAA

FIGURE 96

MAGIPGLLFLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIIYLSSSGDGAQHRDSGSSGKS
RRKRQIYGYSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG
TQKLRVGLFKPKFDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA
LLELKPKPKRKFMKIGVSPPAKQLPGGRIHFSGYDNDPRGNLVRFCDVKDETYDLLYQQCD
AQFGASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW
IKGNYLDCREG

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

FIGURE 97

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCGCCCCACCAGCCATGCTGTTT
CTGGAGCGCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG
GCGTCGACAGCCATCTCTCAATGCGGCCAGGATACCTGTTTCCCCAGCCTGTGGGAAGCCCCA
GCAGCTGAACCGGTTGTGGGCGGCGAGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA
GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC
ACTGCTGCCACTGTTCAGGACAACCTGAACAACCATACCTGTTCTCTGTGCTGCTGGG
GGCCTGGCAGCTGGGGAACCTGGCTCTCGGTCCCAGAAGGTGGGTGTGCTGGGTGGAGC
CCCACCTGTGTATTCTTGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCCTCTCGAG
CGCTCCATACAGTTCTCAGAGCGGGTCTGCCCATCTGCCTACCTGATGCCCTATACCACT
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC
CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTCTATCATCGACTCGGAAGTCTGCAGCCAT
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT
GGAGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG
GCGCTTGGCTGTGGCCGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC
GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA
GCTCCGCGGGCGCGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCGG
CCGCGCGCTCTTACGGGCGAGCGGGACGCGGGGCTCGGATCTGAAAGCGGCGCAGATCCACA
TCTGGATCTGGATCTGCGGCGGCTCGGGCGGTTTCCCCCGCGTAAATAGGCTCATCTACC
TCTACCTCTGGGGGCGGACGGCTGCTGCGGAAAGGAAACCCCTCCCCGACCCGCCGAC
GGCTCAGGCCCCCTCCAAGGCATCAGGCCCGGCCAACGGCTCATGTCCCCGCCCCAC
GACTTCGGCCCCGCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT
AGGTATTGTAAACCTGCCACATATCTTATTTATTCCTCCAATTTCAATAAATTATTTATT
CTCCAAAAA

1000
900
800
700
600
500
400
300
200
100
0

FIGURE 98

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318
><subunit 1 of 1, 317 aa, 1 stop
><MW: 33732, pI: 7.90, NX(S/T): 1
MVVSGAPPALGGGCLGTFSTLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP
WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA
WVEPHPVVSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDG
VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPFITEDMLCAGYLEGERDACLGDSGGPLMC
QVDGAWLLAGIISWGECAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG
SGAAARS
```

Signal sequence:

amino acids 1-32

N-glycosylation site.

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

FIGURE 99

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTA
 CTGCTGGTGGCCACCACAGGCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGT
 GGAGCTGCACAACCTCTACCGGGCCAGGTATCCCCAGCGCCTCAGACATGCTGCACATGA
 GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCACGCACGGCAGTGCGTGTGGGGCCAC
 AACAAAGGAGCGCGGGCGCCGCGGCAGAAATCTGTTCCGCATCACAGACGAGGGCATGGACGT
 GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACTCAGCGCCGCCACCT
 GCAGCCCAGGCCAGATGTGCGGCCACTACACGAGGTGGTATGGGCCAAGACAGAGAGGATC
 GGCTGTGGTTCCCACTCTGTGAGAAGCTCCAGGGTGTGAGGAGACCAACATCGAATTACT
 GGTGTGCAACTATGAGCCTCCGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC
 CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAATCCCTCTGTGAACCCATCGGAAGC
 CCGAAGATGCTCAGGATTTGCCCTTACCTGGTAAC TGAGGCCCATCCTTCCGGGCGACTGA
 AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT
 TGGTAAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAACCCAGGCC
 CCAACTTCTTTAGCAACGAAAGACCCGCGCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC
 AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGATGAGGAGCCAGTTA
 CTTTCCCAAATCGACCCATGTTCTCTATCCCAAATCAGCAGACAAAGTGACAGACAAAACA
 AAAGTGCCCTCTAGGAGCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG
 GGAACTCCTACCCCATGCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCCTCCTTCCAGTG
 AGGTCTTGGCCTCAGTTTTTCCAGCCAGGACAAGCCAGGTGAGCTGCAGGCCCACTGGAC
 CACACGGGGCACACCTCCTCCAAGTCCCTGCCAATTTCCCAATACCTCTGCCACCGCTAA
 TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCGTCTTGGCAGGTGCAGAGGGCCCTGACA
 AGCCTAGCGTTGTGTGAGGCTGAACTCGGGCCCTGGTCAITGTGTTGGGGCCCTCTCCTGGGA
 CTACTGCTCCTGCCTCCTCTGGTGTGGCTGGAATCTCTTGAATGGGATACCACTCAAAGGG
 TGAAGAGGTCAAGCTGCTCCTCTGTCATCTTCCCCACCTGTCCCCAGCCCCATAACAAGATA
 CTTCCTGGTTAAAGGCCCTCCGGAAGGGAAAGGCTACGGGCATGTGCCCTCATCACACCATCC
 ATCCTGGAGGCACAAGGCCCTGGCTGGCTGCGAGCTCAGGAGGCCGCTGAGGACTGCACACC
 GGGCCACACCTCTCTGCCCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT
 CACTGCCTACCTGGCCTGGGGCTGTCTGCCCCACAGCATGTGCGCTCTCCTGAGTGCCCTG
 TGTAGCTGGGGATGGGGATTCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC
 TTTGAGTGGGGAGGCAGGACGAGGGAAGGAAAGTAACCTCTGACTCTCCAATAAAAACCT
 GTCCAACCTGTGAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 100

MHGSCSFLMLLLPLLLLLVATTGVPVGALTDEEKRLMVELHNLRYAQVSPITASDMLHMRWDEE
LAFAKAYARQCVVGHNKERRRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ
MCGHYTQVVWAKTERIGCGSHFCEKLGVEETNIELLCVNYEPPGNVKGKRPYQEGTPCSQC
PSGYHCKNSLCEPIGSPEDAQDLPLYLVTEAPSFRAEASDSRKMGTTPSSLATGIPAFLVTEV
SGSLATKALPAVETQAPTSLATKDPSPMATEAPPCVTTTEVPSILAAHSLPSLDEEPVTFPKS
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEALPPSSEVLAS
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGDPKPSVV
SGLNSGPGHVWGPLLGLLLLPLVLGAGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,
250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

FIGURE 101

GTAACTGAAGTCAGGCTTTTCATTGGGAAGCCCTTCAACAGAATTCGGTCATTCTCCAAGTTATGGTGGACGT
 ACTTCGTGTGTTCTCCCTCTGCTTGCTTTTTCACATTAGCAGACCGGACTTAAGTCACAAACAGATTATCTTTCAT
 CAAGGCCAAGTTCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAACTGAACACAAATGAATTGGAGACCATTC
 AAATCTGGGACCGACTCTCGGGCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTTGTTGAAATCTCCCTGA
 ACATCTGAAAGAGTTTCAGTCCCTTGAACCTTTGGACCTTAGCAGCAACAATATTTTCAGAGCTCCAACTGCATT
 TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA
 TTGGGCCAACACACTCTCTGTGTTAAAGCTGAACAGGAACCGAACTCAGCTATCCCAACCAAGATGTTTAAACT
 GCCCCTCACTGCAACACTCTCGAATTGAACCGGAACAGGATTAAGAACTGTAGATGAGCTGACATTTCCAAGGCTTTGG
 TGCTCTGAAGTCTCTGAAATGCAAAAGAAATGGAGTAACGAAACTTTATGGATGGAGCTTTTGGGGCTGAGCAA
 CATGGAAATTTTGACCTGGGACCAATAACCTTAACAGAGATTACCAAGGCTGGCTTTTACGGCTTCTGTATGCT
 CGAGGAATCTCATCTGAGCTGAGCAATGCCATCAACAGGATCAGCCCTGATGCTCTGGAGTTCGCCAGAGCTCAG
 TGAGCTGGACCTAACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCTTAAGCTTACTAAATAC
 ACTGCACTTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCCTTCGGGGCTTTCCAGTTTAAAGACTTT
 GGATCTGAAGAACCAATGAAATTTCTGGAATATTGAAGCATGAATGGTGCTTTCTCTGGCTTGAAGCATTTGGAGCA
 TCTAGACCTGAGTGACAACGCAATCATGCTTTTACAAGGCAATGCATTTTCAAAAATGAAGAACTGCAACAAT
 GCATTTAAATACATCAAGCCTTTTGTGGATTCAGCTTAAATGGCTTCCACAGTGGGTGGCGGAAACCAATTT
 TCAGAGCTTTGTAAATGCCAGTTTGGCCCTCTCAGCTGCTAAAAGGAAGAACATTTTGTGTTGTATGCCAG
 TGCTTTGTGTGTGATGATTTTCCAAAACCCAGATCAGGTTTCAGGCAGAAACACAGTGGGCAATAAAGGTTCC
 CAATTTGAGTTTCTCTGCTCAGCTGCGCAGCAGCAGTGATTCCCAATGACTTTTGTCTGGAAAAAGCAATGA
 ACTACTGATGATGCTGAAATGGAAAAATATGACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACAC
 CATCTCTCGCTCGCGGGTGGAAATTTGCGAGTGAGGGGAAATATCAGTGTGTCTCATCTCCAATCACTTTGGTTT
 ATCTACTCTGTCTCAAGGCCAGCTTACAGTAAATATGCTTCCCTCATTTACCAAGACCCCATGGATCTCACAT
 CCAGCTTGGGCCATGGCCAGCTTTGGAGTGTCTGCTGTGGGGCACCCAGTCCCGCAGATAGCTGGCAGAGGA
 TGGGGGCACAGACTTCCAGCTGCAAGGGAGAGACGATGCATGTGATGCCGAGGATGACGTGTTCTTTATGCT
 GGATGTGAAGATAGAGGACTTTGGGGTATACAGCTGCAAGCTCAGAACAGTGCAGGAGATGATTTTCAGCAATGC
 AACTCTGACTGCTCTAGAAACCACTCATTTTTCGGGCCACTGTTGGACCGAACTGTAAACCAAGGGAGAAACAGC
 CGTCTCAGTGATGCTTGTGGAGGAAGCCCTCCCTTAACTGAACTGGACCAAGATGATAGCCATTTGGTGGT
 AACCGAGAGGCACTTTTTCGAGCAGGCAATCAGCTTCTGATTTATTTGGAGCTCAGATGTGAGTATGCTGGGAA
 ATACACATGTGAGATGTCTAACACCTTGGCACTGAGAGAGGAAACGTGGCCCTCAGTGTGATCCCACTCCAAC
 CTGCGACTCCCTCAGATGACAGCCCACTCGTTAGACGATGACGGATGGGCCACTGTGGGTGTGCTGATCTATAGC
 CGTGGTTTGTCTGTGTGGTGGGCAGTCACTCGTGTGGTGGTTCATCATATACACACAAGGCGGAGGAATGAAGA
 TTGACGACTTACCAACACAGATGAGACCAACTTGCAGCAGATATTTCTAGTTATTTGTGATCTCAGGGAACGTT
 AGCTGACAGGCGAGTGGGTACGTGCTTTCAGAAAGTGAAGCCACCAAGCTTTGTCACATCTCAGTGTCTGG
 ATTTTTCTTACCAACATGACAGTGTGGGACCTGCCATATTGACAATGACAGTGAAGCTGTATGTGGAACTGCA
 CACAGATCTCTCTTTTGTGCTTTTGGGATCCACAGGCCCTATGTATGACAGATCTCAGTGTCTGG
 TCCTTTTGAACATATCATACAGTTGCACTGCTGACCAGAAACAGTTTAAATGGACCACTATGAGGCCAGTTA
 CATAAAGAAAAAGGAGTGCTTCACTGTTCTCATCTTCAGAGAATCTGCGAAGCCAGTTCAGTAATATTC
 GTGGCCTTCACATGTGAGGAAGCTACTTAACTAGTTACTCTCAATGAAGACCTTGAATGAAATATCTGTG
 TCTAAACAGTCTCTTTAGATTTTATGTCGAAATCCAGAGCCAGCGTCCGTTGCTCCAGTAACTTTCTATGGG
 TACCTTTGAAAAGCTCTCAGGAGCCTCACCTAGATGCTTCAAGCTTTGGACAGCCATCAGATTTGTCAGCC
 AAGAGCCTTTTATTTGAAAGCTCATTTCTTCCCAAGCTTGGACTTGGGTGAGGGAAGATGGGAAAGAGAC
 AGATTTTTCAGGAGAAATCATTTGTACCTTTAAACAGACTTTAGAAAATACAGAGCTCCAAATTTTTCAGTC
 TATGACTTGGACACATAGACTGATGATGAGACCAAGGAAAAAGCTTAACTACTACTCTCAGTGAACCTTTTATTA
 AAAGAGAGAGAAATCTATGTTTTTTAAATGGAGTTATGAATTTTAAAGGATATAAAATGCTTTTATATACAGAT
 GAACCAAAATTAAGAAAGTTATGAAGATTTTATCTGGGAATGAGTCTCATATAAGAACTACCTTTTAACTAT
 TTTTAACTTTGTTTATGCAAAAAGATCTTACGTAAATTAATGATATAAACTATGATTTTATGTATT
 TATTAATGCCAGATTTCTTTTATGAAAATGAGTTACTAAAGCAATTTAAATATACCTGCTGCTGTACCAATTT
 TTAATAGAAGTTACTTCATTATTTTGCACATTATTTTAAATAAAGTGTCAATTTGAA

FIGURE 102

MVDVLLLFSLCLLFHISRDPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVVSAN
ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLNLNRRVTSMEP
GYFDNLANTLLVLKLNRRNRISAIPPKMFKLPLQLHLELNRRNKIKNVDTLTFQGLGALKSLKM
QRNGVTKLMDGAFWGLSNMEILQLDHNNTLITKGLWLYGLMLQLHLHLSQNAINRISPDawe
FCQKLSeldLTFNHLSRLDDSSFLGLSLNTLHIGNNRVSIADCAFRGLSSLKTLDLKNN
ISWTIEDMNGAFSGLDKLRRLILQGNRIRISITKKAFTGLDALEHLDSLDAIMSLQGNAFSQ
MKKLQQLHLNTSSLLCDCLKWLQWVAENNFOQSFVNASCAHPQLLKGRSIFAVSPDGFVCD
DFPKQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKDNELLHDAEMENYAHRAQG
GEVMEYTTILRLREVEFASBGKYQCVISNHFGSSYSVKAKLTVMNLPSTKTPMDLTIRAGA
MARLECAAVGHGAPQIAWQKDGDTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN
SAGSISANATLTLVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPKLNWTKDDSPLVVTER
HFFAAGNQLLIIVDSVDSDAGKYTCESMNTLGTERRGNVRLSVIPTPTCDSPQMTAPSLDDDG
WATVGVVIIAVVCCVVGTSVLVWVVIYHTRRRNEDCSITNTDETNLPAIDPSYLSQGTLD
RQDGYVSSGESGSHHQFVTSSGAGFFLPQHDSSGTCCHIDNSSEADVEAATDLFLCPFLGSTGP
MYLKGNVYGSDFPETYHTGCSPPDPTVLMHDHYEPSYIKKKCYPCSHPSSESCERSFSNISW
PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLDIFSANPEPASVASSNSFMGTFGKALRRPHLDA
YSSFGQPSDCQPRAFYLKAHSSPDLDSSGEEDGKERTFQEEHNHICTFKQTLENYRTPNFQS
YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

FIGURE 103

GGGGAGAGGAATTGACCATGTAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTGGGTGCCTTGCAAAATG
 AAGGATGCAAGGACGAGCTTTCTCTGGAAACCGAAGCAATGGATAAACTGATTGTGCAAGAGGAAGGAAGAAC
 GAAGCTTTTTCTGTGAGCCCTGGATCTTAACACAAATGTGTATGTGTCACAGAGGAGCAATTCAAGAATGAAA
 TAAACAGAGTTAGACCCCGGGGGTGTGGTGTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCCTCC
 CCACCCCCAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAGTATGTTTATTCTC
 TATAAAGGAGAAATGAGCCCAAGGAGATATTTTGGAAATGAAAGTTTGGGGCTTTTATGATAAAGTAAAGAACT
 GGTGGTGGTGGTGTTCCTTTCTTTTGAATTTCCCAAGAGGAGAGGAATTAATAATACATCTGCAAAAGAAA
 TTTGAGAGAAAGAAAGTTGACCGCGGAGATTGAGGCATTTGATGGGGGAGAGAAACGAGCAGCAAGTTGGA
 TTTGTGCTATGTTGACTAAAAATTGACGGATAATTGCAAGTTGGAATTTTCTCATCAACCTCTTTTTTTAAAT
 TTTTATTCTTTTGGTATCAAGATCATGCGTTTCTCTGTCTTAACACCTGGATTTCATCTGGATGTGTCT
 GTGATCAGTCTGAAATACAACTGTTTGAATTCAGAAGGACCAACACAGATAAATATGATGTTGAAACAAAGAT
 GACCTTACATCCACAGCAGATAATGATAGGCTCTAGGTTTAAACAGGGCCCTATTGACCCCTGCTTGTGGTGTCT
 GCTGGCTCTTCAACTTCTGTGGTGGCTGGTCTGGTGGCGGCTCAGACCTGCCCTCTGTGTGTCTCTGCAGCA
 CCAGTTGAGCAAGGTGATTGTGTGCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT
 GCTGAACCTCCATGAGAACCAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCCTTGGAAATCCT
 ACAGTTGAGTAGGAACCATATCAGAACCATGAAATGGGGCTTTCATGGTCTGGCGAACCTCAACACTCTGGA
 ACTCTTGTACAATCGTCTTATACCACTCCCGAATGGAGCTTTTGTATCTTGTCTTAACTGAAGGAGCTCTGGTT
 GCGAAACACCCCATGAAAGCATCCCTCTCTATGCTTTTAAACAGAACTCTCTTCTTGGCGGCTAGAGCTTAGG
 GGAATGAAAGACTTTCATACATCTCAGAAGTGCCTTTGAAGGTCTGTCCAACCTGAGGTATTTTGAACTTGC
 CATGTGCAACCTTCGGGAAATCCCTAACCTCACAACCGCTCATAAACTAGTACGCTCTTTTGGGAATCA
 TTTATCTGCCATCAGGCTGGCTTTTCCAGGTTTGTGATGCATCTTCAAAACTGTGGATGATACAGTCCCAGAT
 TCAAGTGTATGAACGGAATGCTTTTGACAACCTTCAGTCACTAGTGGAGATCACTGGCACACAATAATCTAAC
 ATTAGTGCTTCACTGACCTCTTCACTCCCTTGATCATCTAGAGCGGATACATTTACATCAACACCTTGGAACTG
 TAACTGTGACATACTGTGCTCAGCTGTGGTATGATAAAAGACATGGCCCTCGAACACAGCTTGTGTGGCGGGT
 TAACACTCTCCCAATCTAAAGGGGAGGTACATTTGGAGAGCTCGACAGAAATACTTCACATGCTATGCTCCGGT
 GATTGTGAGAGCCCTGACAGCTCAATGTCACTGAAGGATGGCAGCTGAGGTGAAATGTGCGGGCTCCACATC
 CTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCTGACACATGGGGCGTACAAAGTGGCGGATAGCTGT
 GCTCAGTGATGGTACGTTAAATTTCACAAAATGTAAGTGTGCAAGATACAGGCTGTACACATGTATGGTGTGATAA
 TCCGTTGGGAATACTACTGCTTCAGCCACCTGAAATGTTACTGCGACCAACCACTACTCCTTTCTCTACTTTTTC
 AACCGTCAAGTAGAGACTATGGAACCGTCTCAGGATGAGGCAACGACACAGATAACAAATGTGGGTCCCCTCC
 AGTGTGCTGCTGGGAGACACCAATGTGACCACTCTCTCACACACAGAGCAAGGTCGACAGAGAAAACTT
 CACCATCCAGTGAATGATATAAACAGTGGGATCCAGGAATGTGAGGTATGAAGAATCAACAAATCATCAT
 TGGGTGTTTGTGGCCATCACACTCATGCTGAGTGATGCTGGTCACTTTTCTACAAGATGAGGAAGCAGCAACA
 TCGGCAAAACCATCACGCCCCAACAGGACTGTGAAATTTAATGTGGATGATGAGATTACGGGAGACACACC
 CATGGAAGCCACCTGCCCATGCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT
 CAACCAACACAAACAGCTTAAACAAATAAATTCATAACAGCTTCAGTGACAGACCGTTATTGATCCGAATGAA
 CTCTAAAGCAATGTACAAGAGACTCAAATCTAAACATTTACAGAGTTTCAAAAAACAAACATCAAAAAAAA
 GACAGTTTATTAATAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAACAA
 AAAAGAAAAGAAATTTATTTATTAATAATCTATTTGTGATCTAAAGCAGACAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 104

MLNKM TLHPQQIMIGPRFNRLFDPLLVLLALQLLVVAGLVRAQTCPSVCSNQSFKVIC
VRKNLREVPDGI STNTRLNLHENQIQIIKVNSFKHLRHL EILQLSRNHIRTIEIGAFNGLA
NLNLTLELFDNRLTTIPNGAFVYLSK LKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNRLRYLNAMCNLREIPNL TPLIKLDEL DLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAPDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRATSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVNSVGN
TTASATLNVTAAATTPFSYFSTVTVETMEPSQDEARTDNNVGPTPVVDWETTNVITSLTPQ
STRSTKFTFTIPVDINSGIPGIDEVMKTKIIIGCFVAITLMAAVMLVIFYKMRKQHRQN
HHAPTRTVEI INVDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,
442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

FIGURE 105

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGCA
TGAAAGGGGCTTCGCGCGCCGGAGTAAAGAAAGGAAATTGACCGGGCAGCGGAGGGAGAGCGCGCACCGCAGCCG
GAGGGCGGGCTGCACCTCTCGCTGCGAAGTTTGTGCGGGGCCGAGCGCGCGCGCTTCGGGAGTCTTCGGGTAGA
GACCTAGGCGCGCTGACCCGCGTAGCGCGCGCGGAGCCTTCGTCGCGCGCGCGCGGGGTTGGGGCTGCTGCTGTGC
CGCGTGTGGGGCGCGCTGCGCGGTTCGACAGCGCGCGCTGCGGGGAACTCGGCGAGCCCTTCGCGGTAGCCGCG
GAGCGCCCATGCGCCACTACTTGCCTGCGCTGCGGGGACCTGCTGGAGCTCAGTCGTAAGCGGCTAGCGCGCTCT
CCCGAGGCACTCCGCTCGCTCGCTGCGCTGCGCTTAAGTCAACAAGATATCTTTTCATCAAGGCAAGTCTC
ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAGTGAACAAACATGAATTGGAGACCATCCAAATCTGGGACCA
GTCTCGGCAAAATGACCTTCTCTTCCTTGGCTGGAAACAGGATTGTTGAATACCTTCGACATCTCGGAGCTTGAAGAG
TTTCAGTCCCTTGAACCTTTGGAACCTTAGCAGCAACAATATTTAGAGCTCCAACTGCAATTCAGCGCTTACAG
CTCAAAATATCTGATTTCAACAGCAACCGAGTCACTCAATGGAACCTCGGATATTTGACAAATTTGGCCAACCA
CTCCTTGTGTAAAGCTGAACAGGAACCGAATCTCAGTATCCCAACCAAGATGTTTAACTGCGCCCACTGCA
CATCTCGAATGAACCGAAACAAAGATTAAAAATGTAGATGGACTGACATTTCCAAGGCTTGGTGCTCTGAAGTCT
CTGAAAAATGCAAGAAATGGAGTAACGAACTTATGGATGGAGCTTTTGGGGGCTGAGCAACATGGAATTTT
CAGCTGGAGCATAACCTTAAACAGAGATTACCAAGGCTGGCTTTACGGCTTGTCTGATGCTGCGAGAACITCAT
CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCTCGGGAGTTCTGCGAGAAGCTCAGTGAGCTGGACCTA
ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCTTAAGCTTACTAAATACATGCACATTTGGG
AACCAACAGAGTCACTACATTTGCTGATTTGCGCTTCCGGGGCTTTCAGATTTAAAGACTTTGGATCTGAAGAAC
AATGAATTTCTTGAGACTTGTGAAGACATGAATGGTGTCTTCTTGGCTTGACAACTGAGGAGCTGATCTC
CAAGGAATCGGATCCGCTCTATTACTAAAAAGCCTTCAGCTGTTGGAGTCAATTTGAGCATCTAGACCTGAGT
GACAAACGAATCACTCTTTCAAGGCCAATGCATTTTCAAAATGAAGAACTGCAACAAITTTGCAATTTAATACA
TCAAGCTTTTGTGCGATTGCGCACTAAAATGGCTTCCACAGTGGGTGGGAGGAGAACTTTCAGAGCTTTGT
ATATGCAATTTTGCGCCACTCTCAGCTGCTAAAAGGAAGAGCAATTTTGTGCTTTAGCCCGAGTGGCTTTGTGT
GATGATTTTCCAAACCCAGATACCGGTTTCAAGCCAGAAACAGCTCGGCAATAAAGGCTTCAATTTGAGTTTC
ATCTGCTCAGCTGCCAGCAGCAGTGATTTCCCAATGACTTTTGTCTTGGAAAAAGACAAATGAACACTCTGCATGAT
GCTGAAATGAAAAATTTATGACACTCCGCTCGGGGCCAAGGTGGCGAGGTGATGGATGATACACCATCTTGGCTG
CGCAGGTGGAATTTGCGAATGAGGGGAAATATCAGTGTGTCATCTCAATCACTTTGGTTTCACTCACTCTGTC
AAAGCCAAAGCTTACAGTAATATGCTTCCCTCATTCACCAAGACCCCAATGGATCTCACCATTCCGAGCTGGGGC
ATGCGACGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCAGATAGCTGCGCAGAGGATGGGGGCGACAGAC
TTCCAGCTGCAAGGGAGAGCAGCATGCATGTGATGCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGATA
GAGGACATTTGGGTATACAGCTGCACAGCTCAGAACAGTCAGGAAATATTTAGCAAAATGCAACTCTGATGTC
CTAGAAACCACTCATTTTTTGGGGCACTGTTGGACCGAATCTGAACCAAGGAGAAACAGCGCTCTACAGTGC
ATTGCTGGAGGAAGCCCTCCCCATAAATGAACTGGAACAAAGATGATAGCCCATTTGTTGTAACCGAGAGGCA
TTTTTGTGACGAGCAATCAGCTTCTGATTTATTTGGACTCAGATGTCAAGTATGCTGGGAAATACATCTGTGAG
ATGCTAACACCTTTGCACTGAGAGAGGAAACGTGCGCTCAGTGTGATCCCCACCTCAACCTGCGCATCCCCCT
CAGATGACAGCCCACTGTTAGACAGTGAAGGATGGGCTGCTGGGTGTCTGTGATCATAGCCGTGGTTCTTCGAT
GTGGTGGGCACTGCTCACTCGTGTGGGTGGTCACTATATACCAACAAAGCGGAGGAAATGAAGATTGAGCATATACC
AACACAGATGAGACCACTCTGACAGCAGATATCTTAGTTATTTGTCACTCAGGGAACTGTTAGCAGCAGCAG
GATGGGTACGTGCTTTCAGAAAGTGGAAAGCCACCAACAGTTTGTCACTCTCAGTGCTGGATTTTCTTACCA
CAACATGACAGTAGTGGGACTGCCATATTGACAAATGACAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTT
CTTTGTCCGTTTTTGGGATCCACAGGCCCTATGATTTTGAAGGGAATGTGTATGGCTCAGATCCTTTTGAACA
TATCATACAGGTTGCGACTCTGACCCCAAGAACGTTTAAATGGACCACTAGAGCCAGTTTACATAAAGAAAAAG
GAGTGTACCCATGTTCTACATCTTTCAGAGAATCTCTGCGAACGGAGCTTCAGTAATATATCTGTGCGCTTCACT
GTGAGGAGCTACTTATCTGCTTTACTCTCAATGAAGGACCTTGAATGAAAAATCTGTGCTCAACCAAGCTTC
CTTTAGATTTTATGCAAACTCAGAGCCAGCGTGGCTGCTCAGTAATTTCTTTCAGGGTACCTTTTGAAGAA
GCTCTCAGGAGACCTCAGCTAGAGCTGCTTCAAGCTTTGGACAGCAGCATCAGATGTGACGCAAGAGCCTTTAT
TTGAAGCTCATTTCTTCCCAGACTTGGACTCTGGGTGAGAGGAAGTGGGAAGAAAGGAGCAGATTTTCAAGAA
GAAATCACTTTGACCTTTTAAACAGACTTTAGAAAACTACAGGACTCAGAAATTTTCACTTTATGACTTGGC
ACATAGCTGAATGAGACCAAGGAAAGCTTAACTACTACCTCAAGTGAATTTTATTTAAAGAGAGAGAAAT
CTTATGTTTTTAAATGAGGTATGAATTTTAAAGGATAAAATGCTTTATTTATACAGATGAACAAATATAC
AAAAAGTTATGAAGAAATTTTATACCTGGGAATGATGCTCATATAAGAAATACCTTTTAACTATTTTAACTTTG
TTTTATGCAAAAAAGTACTACGTAAATTAATGATATAAATCATGATATTTTATGATATTTTAAATGGCAGA
TTTCTTTTATGGAAGATGAGTACTAAAGCAATTTTAAATTAATACCTGCTTGTACCATTTTAAATAGAGTT
ACTTCATTATATTTTGCACATATATTTTAAATAAATGTGCAATTTGAAAAAAGAAAAAAAAAAAAAAAAAAAAA

FIGURE 106

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGRGELGQPSGVAAERPCPTTCRCLGDLDDCSR
KRLARLPEPLPSWARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT
LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFFPALQLKYLYLNSNRVTSMEPGY
FDNLANTLLVLKLNRRISAI PPKMFKLQQLHLELNRNKIKNVVDGLTFQGLGALKSLKMQR
NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDWEFC
QKLSELDLTNNHLSRLDDSSFLGLSLNLT LHIGNNRVSYIADCAFRGLSSSLKTLDLKNNETS
WTIEDMNGAFSGLDKLRRLILQGNRIRISITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK
KLQQLHLNTSSLLCDCQLKWLPPQWVAENNFQSFVNASCAPHQLLKGRSIFAVSPDGFVCDFF
PKQITVQPPTQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENYAHLRAQGGGE
VMEYTTILRLREVFEASEGKYQCVISNHFSSSYSVKAKLTVMNLPSTFTKTPMDLTIRAGAMA
RLECAAVGH PAPQIAWQKDGDTFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSTAQNSA
GSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF
FAAGNQLLLIIVDSVDSDAGKYTCEMSNTLTGERGNVRLSVIPTPTCDSPQMTAPSLDDDDGWA
TVGVVIAVVCVVGTSLVWVVI IYHTRRRNEDCSITNTDETNLPAIDIPSYLSSQGTLDLRQ
DGYVSSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPLFGSTGPMY
LKGNVYSDPFETYHTGCSPPDRTVLMDHYEPSYIKKKECYPCSHPSSESCERSFSNISWPS
HVRKLLNTSYSHNEGPGMKNLCLNKS SLD F SANPEPASVASSNSFMGTFGKALRRPHLDAYS
SFGQPSDCQPRAFYLKAHSSPDLDGSGSEEDGKERTDFQEENHICTFKQTLNRYRTPNFQSYDLDT

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 808-828

N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,
1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,
513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

FIGURE 107

CAAAACTTGCCTGCGGAGAGCGCCAGCTTGACTTGAATGGAAGGAGCCGAGCCCGCGGAGCGCAGCTGAGAC
 TGGGGGAGCGCGTTTCCGCTTGTGGGGCGCGCTCGCGCGCGGGGCGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC
 CTGCTCCACGAGCGCCACTTGGTGTGAACCGGGAGAGCCCTGGGTGGTCCCGTCCCTTACCTCTTTATATA
 GAAACCTTCCACATCTGGGAAGGCAGCGCGAGGAGGAGGCTCATGGTGAGCAAGGAGGCGCGCTGATCTGAG
 GCGCACAGCATTCGAGTTTACAGATTTTACAGATACCAAATGCGAAGCGAGGAGGAGCAACAGCTTGCCTGGT
 TCCATCAGCCCTGGCGCCAGCGCATCTGACTCGGCACCCCTCGAGGACCAATGGCCAGAGCGCGGTGTCTGC
 TGCTCTGTCTGTCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGGCGTGAAGGCGCCAGGATTTGGCCGAA
 GTGGCGGCGACAGCTGAGCCCCAGAGAAACGAATTTGGCGAGGAGGAGCGGTGTGGTACTGAGCCCTGAGG
 AGCCCGGCGCTGGCCAGCGCGGTGACGTGCGCCCGAGACTGTGCTGTTCCTGAGGCGCGCTGTGGACTGTG
 CGGATTTGACTTGCCTGAGTTCCTCGGGGACCTGCCTGAGCACCAACCACTTATCTCTCAGAAACCAACGAG
 TGGAAAGATTACTACCTTGAGGAGCTTCTCCCGGTGACACCGCTGGAGACTGAGACCTGCAAAACCAACCGCTGAG
 CTTCCGAGGGCTCCAGAGAAGGCGTTTGAAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAAATAACAAGC
 TGACCTTTGGCACCCCGCTTCTTCCCAAACGCCCTGATCAGTGTGGACTTGTCTGCCAACTATCTACCAAGATCT
 ATGGGCTCACCTTTGGCCAGAGCGCAAACTTGAGGTCTGTGTACTTGCACAAACAACAGCTGGCAGACGCGGGG
 TGCCGAGCAACATGTTCAACGGCTCCAGCAACGCTCGAGGTCTCATCTGTCCAGCACTTCTTGGCCACGCTGC
 CCAAGCACCTGCGCGCTGCCCTGTACAAAGCTGCACCTCAAGAACAAACAGCTGGAGAAGATCCCCCGGGGCGCT
 TCAGCGAGCTGAGCAGCTTGCAGCTTGTGAGAAGAAGCCATCTGCAGAACTACCTGACTGACGAGGGCTGGACAACGAGA
 CTTCTTGGAACTCTCCAGCTTGAAGTACTTGGATCTGTCCAGCAACAACTGTCTCGGGTCCAGCTGGGCTGC
 CGCGCAGCTGGTGTCTGCTGCACTTGTGAGAAGAAGCCATCCGAGCGCTGGACCGCAATGTGTGACGCCCATCC
 GCAGCTGGAGTACTTGTGCTGCACTTGCACAGCAACAGCTGCGGGAGCGAGGCACTCCCAAGCTTCCAGGGCC
 TCAAGCGGTGACCAACGCTGCACTGTACAAACAGCGCTGAGGCGCGTGCACAGTGGCTTGCCTGCGCGGTGC
 GCACCTCATGATCTCTGCAACCAACAGCTCACAGGCTTGGCGCGAAGACTTTCGCCAACCTACTTCTGTGAGG
 AGCTCAACTGATGACTACAACCGCATCACAGCCACAGGCTGCACGCGCGAGCGCTTCCGCAAGCTGGCGCTGTCTG
 GCTCGTGGACCTTGTGGGCAACCGGCTGCACACGCTGCCAAGCTGGGCTGTCTCGAAATGTCCATGTGCTGAAGG
 TCAAGCGCAATGAGCTGGCTGCTTTGGCACGAGGCGCGTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACTCA
 CCGACCAACAGCTTGCAGCGGAGCTTGGGCGCCGCTGCTGGGTGAGACTTGCCTCATCTGCGAGCTGCTGGACA
 TCGCCGCGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCGGAGTCACTTGAGTACTGTACTCTGAGAACAAACA
 AGATTAGTGGGTGGGCGGCGCAATGCTTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAAACAAGC
 TGGCTGTGGGCTCGGTGGTGGACAGTGCTTCCGAGGCTGGAAGCACTCGAGGTCTTGGACATTTGAAGGCAACT
 TAGAGTTTGGTGACATTTCCAAGGACCGTGGCGCTTGGGGAAGGAAAGGAGGAGGAGGAGGAGGAGGAGG
 AGGAAGAGGAAACAAATAGTAGCAAGGTGATGAGATGTGACCTAGGATGATGAGACGCGCGGACTCTTTCTGCG
 AGCACACGCTGTGTGTGTGAGCGCCCACTCTGCGGTGCTCACACAGACACACCCAGCTGCACATGAGGCA
 TCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCACGCGGTGTCCACGCGCAGACATGCG
 ACACACATCACACCTCAAACACCCAGCTCAGCCACACACAACTACCTCCAAACACCAACAGTCTCTGTCTCAC
 CCCCATCCCGCTGCGACGCGCTCTGAATCATGACAGGGAAGGCTCTGCCCTGCCCTGGCACACAGGCAACCA
 TTCCCTCCCCCTGTGACATGTGTATGCGTATGCATACACACCAACACACATGCAACAGTCACTGTGCGAA
 CAGCCCTCAAAGGCTATGCCACAGACGCTTGTGCCAGCCAGGATCAGGCTACAGCTCGCGCTGTCTGCCCT
 GTCCATCTGTCCGTCCGTCCAGGTGCTGGGCACTCATCTGCTGAAGTCCCTCTGCGCACGCGCTGGCAGGACA
 CAGGCACTTTTCCAATGGCAAGCCAGTGGAGGAGGATGGGAGAGCCCTTGGTGTCTGTGGGCTTGGGG
 CAGGATGAGCAGAGGTGATGGGGCTGGGCTGAGCCAGGAGGAGGACCCAGCTGCACATGAGGAGACCTTTT
 GTTCTTCAAGCCTGTGGGGGAAGTTCGGGGTGCCTTTATTTTTTATTTCTTTCTAAGGAAAAAATGATAAAAT
 CTCAAAGCTGATTTTTCTTGTATAGAAAACTAATATAAAGCATTATCCCTATCCCTGCAAAAAA

FIGURE 108

MEGEEAEQPAWFHQPWPRGASDSAPPAGTMAQSRVLLLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPREENEFABEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFFPGDLP
EHTNHLSSLQNNQLEKIYPEELSRHLRLETLNLQNNRLTSRGLPEKAFEHLETLNLYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDMMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKHLKNNKLEKIPPGAFSELSSRLRELYLQNNYLTDEGLDN
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSDANVLTPIRSLEYLLLSHN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYF
LEELNLSYNRITSPOVHRDAFRKLRLSLDLSGNRLHTLPPGLPRNVHVLKVKRNEAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKD
RGRLGKEKEEEEEEEEEEEETR

Signal sequence:

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,
477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,
535-557

FIGURE 109

GGGAGGGGGCTCGGGCGCGCGCAGCAGACCTGCTCCGGCGCGCGCCTCGCCGCTGTCTCTCGGGAGCGGCGAG
CAGTAGCCCGGGCGCGGAGGGCTGGGGGTTCTCGAGAGTCTCAGAGGGGGCGCTCCATCGGGCGCCACACCC
CAACCTTTCTCTCGCGCGCGCTGCGCTGCGCCCCAGGACCGCTGCCAACATGGATTTTCTCTGGCGCTGGT
GCTGGTATCTCGCTCTACTCGAGCGCGCGCCGAGTTGCAGCGGAGGTGGCCAGGCAAAATAGTGTCACTCGAT
TGGCCTATGTCTTATGGTGGGAGATGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTGCAGCC
TGTGTGCCAACACAGATGCAAACTGGTGAATGTATCGGGCCAAACAAGTGCAGTGTCACTCTGGTTATGCTGG
AAAAACCTGTAACTCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGAAGCACAGGTGCATGAACACTTA
CGGCAGCTCAAGTGCTACTGTCTCAACGAGATATGCTCATGCCGGATGGTTCTGTCTCAAGTGCCCTGCACCTG
CTCCATTGGCAAACTGTCAAGTGTGGTGTGATGTTGTTAAAGGACAAATACGGTGGCAGTGCCCATCCCTGGCCCT
GCACTGGCTCTGATGGGAGGACCTGTGTAGTGTGTGAATGTGTCTACAGGAAGAGCCCTCTGCCCTAGATT
TAGGCAATGTCTCAACACTTTGGGAGCTACATCTGCAGTGTCTAAAGCTCTCGATCTCAGTATTTGGAG
CAAATATCAATGTCAATGATAGACGAATGCTCAGTTGGTCAGTATCAGTGCAGCAGCTTTGTCTCGATGTTATAA
CGTACGTGGGTCTTACAAGTGCAAATGTAAAGAAGGATACAGGGTGTATGGACTGATCTTGTGTGTAATCCCAAA
AGTTATGATTGAACCTCAGGTCCAATTCAATGTACCAAAGGAAATGATACCATTTTAAAGGGTGACACAGGAAA
TAATAATTGGATTCTCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATCTCTCTATCATACCAA
CAGSCCTACTCTTAAGCGAACAAACAGACCTACACCAAAGCCACACCAATTCTACTCCACACCCACACACC
CTGCGCAACAGAGCTCAGAACCTCTTACCACCTACAAACCCGAAAGGCCAACACCCGGAATGACAACTATAGC
ACCACTGCGCACTACACTCTCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCTCAGAAAACCGAGG
AGATGTGTTCACTGTTCTGTGTACACAGTTGTAATTTTGACATGGACTTTGTGGATGGATCAGGAGGAAAGACAA
TGACTTTGACTGGGAACCAATCAGGGACCCAGCAGGTGGACAAATCTGCAGAGTGTGGGACGCAAGGCCGAGG
GGGAAAAGCTGCACGCTTGGTGTACTCTCTCGGCGCCTCATGCATTCAGGGGACCTGTGCTGTCTATTCAAGCA
CAAGGTGACGCGGCTGCATCTTGGCACACTCCAGGTGTTTGTGAGAAAACAGGTCGCCACGGAGCAGCCCTGTG
GGGAGAAATGTTGGCCATGGCTGGAGGCAAAACAGATCACTTGCAGGGGGCTGACATCAAGAGCGAATCACA
AAGATGATTAAGGGTTGGAAAAAAGATCTATGATGGAAAAATTAAGGAACTGGGATTTATGAGCTTGGAGAG
AGAAGACTGAGGGGCAACCATTTGATGGTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACAGCTG
TTCTCCATATGCACTAAGAATAGAACAGAGGAACTGGCTTAGACTAGAGTATAAGGGGACATTTCTTGGCAGG
GGCCATTGTTGAATACTCTATAAAAAAAGAGGTGTGAAATCTCAGTATCTCTCTCTCTTCTTAAAAAATTAGA
TAAAAAATTTGCTATTTAAAGATGGTTAAAGATGTTCTTACCAAGGAAAGGTAACAAAATATAGAAATTTCCAAA
AGATGTTTGTACTACTAGTAGTATGCAAGTAAAACTTTAGAACTAAATAATTTGACAAAGGCTTAATTTAGG
CATTTCCCTCTTGACCTCTTAATGGAGAGGATTAAGAGGGGAAGAGCCCACAAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCTAGCAGTATTAAGAAAAAAGGAAACTATTTATTTCCAAATGAGAGTATGATGGAC
AGATATTTAGTATCTCAGTAATGTCTAGTGTGGCGTGGTTTCAATGTTTCTTCAATGGTAAAGGTATAAGCC
TTTCAATTTGTTCAATGAGTATGTTTCAAGATTTTTTTTTTTTTTAAAGAGTCTCTCAAGGAAACAGATTCAAGAGC
ATTTTCATCGGGTGCACTCTCTGCTCTGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCTCGCCCC
ACACGGCGAGCCTTCTCTCACTCATCAGTATGATTCAGTTTCTCTTATCAATTTGACATCTCCACAGGTTCAC
AGAACAGTAAATTTTTTGAACAATAGGTACAATAGAAGGTCTTGTGATTTAACTGGTAAAGGAGGAGGCTGG
AGGGGGAATAAATCAATTAAGCCTTTGAGTAACGGCAGAAATATAGGTGTAGATCCATTTTAAATGGTTCAAT
TCTTTTATGTCATATAACTGCAACGCTGAAGATGAAAGGGGAAATAAATGAAATTTTACTTTTCGATGCCAA
TGATACATTCGCACTAACTGATGGAAGAAGTTATCCAAGTACTGTATAACATCTGTGTTTATTTATTTATGTTT
CTAAAAATAAAATGTTAGTGGTTTTCAAATGGCTAAATAAAAAACAATATTGTGAAATAAAAAACACGTGTAGTAT

FIGURE 110

MDFLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCGWARQSWGQCQPVQC
RCKHGEICIPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD
GSCSSALTCSMANCQYGCDEVVKGQIRQCQPSPLHLAPDGRCTVDVDECATGRASCPRFRQC
VNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKKEGYQG
DGLTCVYIPKVMIEPSGPIHVPKNGTILKGD TGNNNWI PDVGSTWWPKTPYIPPIITNRP
TSKPTTRPTPKPTPIPTPPPPPLPTLRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDN
RVQTDPOKPRGDVFSVLVHSCNFDHGLCGWIREKDNLDHWEPIRDPAGGQYLTVSAAKAPGG
KAARLVLPGLRLMHSGLCLSFRRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI
TLRGADIKSESQR

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,
421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTGCCCCCTTtagattGTGA
AATGTGGCTCAAGTCTTCACAACTTTCTTTCTTTGCAACAGGTGCTTGCTCGGGGCTGA
AGGTGACAGTGCCATCACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTC
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA
CACAAATGCCAAATACTTACTTGGGCTCTGTGAATAAGTCTGTGGTTCTGACTTGGAATACC
AACACAAGTTCCACATGATGCCACCCAATGCACTCTGCTTATCAACCCACATGCACTTCCCT
GATGAAGGCAATTACATCGTGAAGTCAACATTGAGGAAATGGAACCTATCTGCCAGTCA
GAAGATCAAGTCAAGTTGATGATCCTGTCAAAAGCCAGTGGTGAGATTCACTCCTCCTCT
CTGGGCTGTGGAGTATGTGGGAACATGACCCTGACATGCCATGTGGAAAGGGGCACTCGG
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC
TCCCCAAAACAATACCTTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT
GCCTGGTGAGGAACCTGTGAGTGAATGGAAGTGATATCATTATGCCCATCATATATTAT
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAAGTAGGGGAAGTGTACTGT
TGACCTTGGAGAGGCCATCCTATTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACT
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAAGCATGGGCTCGCTTAGAAGTT
GCATCTGAGAAAGTAGCCCAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC
CGGCAGGCAAGATGAAACTCATTTACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG
CACAAGAAAGAAAATCATTGTCACTTTAGCAAGTATAACTGGAATATCACTATTTTGTATT
ATATCCATGTGTTCTCTCTCTATGAAAAAATATCAACCCCTACAAAGTTATAAAACAGAA
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAAACATTTTCAGGCCATGAAAGATG
CTCTGGATGACTTCGGAAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG
ATTCCAAGCAGGTCTGTTCAGCCTCTGATTGTGTATCGGGCAAGATTGCAAGTACAGT
GTATGAAGTTATTAGCACATCCCTGCCAGCAGCAAGACCATCCAGAGTGAACCTTTCATGG
GCTAAACAGTACATTGAGTGAAATTCGAAGAAACATTTTAAGGAAAAACAGTGGAAGAAAT
ATATTAATCTGGAATCAGTGAAGAAACAGGACCAACACCTCTTACTCATTATTCTCTTACA
TGCAGAAATAGAGGCATTATGCAAAATGAACTGCAGGTTTTTCAGCATATACACAATGTCTT
GTGCAACAGAAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG
GGAGAACGAAAGTGACAGGGGTTTTCTCATAAGTTTTGTATGAATAATCTCTACAAACCTCA
ATTAGTTCTACTCTACACTTTCACATATCATCAACACTGAGACTATCCTGTCTCACCTACAAA
TGTGGAACCTTTACATTGTTTTCGATTTTTTCAGCAGACTTTGTTTTATTAAATTTTATTATG
TTAAGAAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTGTATTATTGTACAA
CAAAGTAATAAGATGGTTGTCAAAAAACAAACTATGCCTTCTCTTTTTTCAATCACC
AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA
TTTTTTTCAGGAAAGATGGATTCAAATAAATATTCTGTTTTTGCTTTTAAAAA

FIGURE 112

MWLKVFTTFLSFATGACSGLKVTVPSTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLKSGSVNKSVPDLEYQHKFTMMPPNASLLINPLQFPDEGNIVKVNIQNGTLSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTM DYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRVSPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208,
276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237,
239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 113

GCAAGCGCGGAAATGCGCGCCCTCCGGGAGTCTTGCAAGTCCCCTGGCAGTCTGTGGTGTGT
GCTTTGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTTCGCGTCATCACGGACGAGA
ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCTGTCT
TGTCAAAATCTTCAACCGGAATGGGAAAGTTTGTCTGAATGGGAGAGAAGATCTTGAGGTTAA
TATTGCGAAAGTAGATGTCAAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCCTC
TTCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGTCACAGGACTAAG
AAGGACTTCATAAACTTTAAGTGATAAAGAGTGGAGAGATTTGAGCCCGTTTCATCATG
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAAGCTCTTTCAGCTATCTATGTGGA
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT
GTTTTTGTCTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC
AGATTGCCCTTGTCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCCTTCAAAAAAAT
TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA
GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGGAACAACAAGACTTTCCACAGAATGC
CATAAGACAACGCTCTGGGTCCATCATTGGCCACAGATAAATCC**TAGTT**AAATTTTATAG
TTATCTTAATATTATGATTTTGATAAAAAACAGAAGATTGATCATTTTGTGGTTGAAGTG
AACTGTGACTTTTTTGAATATTGCAGGGTTCACTCTAGATTGTCAATTAATGAAGAGTCTA
CATTCAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG
GTTTAAATAGTTCTCTAAATTTTGA AAAATCGTGCCCAAGCAATAAGATTATATGATATTTGT
TTAATAATAACCTATTTCAAGTCTGAGTTTGA AAAATTTACATTTCCCAAGTATTGCATTAT
TGAGGTATTTAAGAAGATTATTTAGAGAAAAATATTTCTCATTGTATATAATTTTCTCTG
TTTCACTGTGTGAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATGTCTCAAG
AAATGTGTATTTCACTGACAATTTCTGTGGTCTTTTAGAGGTATATTTCAAAAATTTCCCTGT
ATTTTTAGGTTATGCAACTAATAAAACACCTTACATTAATTAATTACAGTTTTCTACACA
TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTGAAGTTCATGGTATTTCTCTGTATTC
CAACAAAGTTTGATTTTCTCTGTATTTTCTTACTTACTATGGGTACATTTTATTATTTT
CAAATTTGGATGATAATTTCTTGAAACATTTTTATGTTTGTAGTAAACAGTATTTTTTGT
GTTTCAAACGAAGTTTACTGAGAGATCCATCAAATGAACAATCTGTTGTAATTTAAAAAT
TTGGCCACTTTTTTCAGATTTTACATCATTTCTTGCTGAACCTTCAACTTGAAATGTTTTTTT
TTTCTTTTTGGATGTGAAGGTGAACATTTCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA
TTTTACATTTTGAAATTCAAAGAAGCTTAATATAAAAGTTTGCACTTACTCAGGAAAAAG
CATCTTCTGTATATGTCTTAAATGTATTTTTGTCTCATATACAGAAAGTTCTTAATTGAT
TTTACAGTCTGTAATGCTTGATGTTTTAAAAATAATAACATTTTTATATTTTTTAAAAACAA
ACTTCATATATCTCTGTGTTCTTTCTGTACTGGTAATATTGTGTGGGATTTACAGGTA AAA
GTCAGTAGGATGGAACATTTTAGTGATTTTTACTCCTTAAAGAGCTAGAAATACATAGTTTT
CACCTTAAAGAAGGGGGA AAAATCATAAATACAATGAATCAACTGACCATCTAGTAGAGAC
AATTTCTGTAATGTCCCTTCTTTCTAGGCTCTGTTGCTGTGTAATCCATTAGATTTACAG
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAATATTGTACCAAT
AAAGAGTTTTGGATGTGTAACTTGTGATGCTTTAGAAAAATATCTTAAGCACAAAATAAACCT
TTCTAACCACTTCATTAAAGCTGAAAAA AAAAAAAAAA

FIGURE 114

MAPSGSLAVPLAVLVLLWLGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL
QPEWESFAEWGEDLEVNIKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI
NFI SDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLG LFPVWGSYTVFAL
ATLFSGLLLGLCMIFVADCLCP SKRRRPQYPYPYPSKLLSESAQPLKKVEEEQEAEDEEDVSE
EEAESKEGTNKDFPQNAIRQRSLGPSLATDKS

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCCCTT
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAAGAGCGGG
AAGGTCCTGGCCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAAACTCTGGGTGTCTGC
ATTGCTGATGGCCTGGTTTGGTGTCTGAGCTGTGTGCAGGCCGAATTCTTCACTCTATTG
GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAAGAGCTGGGCCAACAAATGGAAAGCCTTGAC
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAACTGG
TGAAGCGGCTTAAACACAGACTGGCTGCGCTGGAGGACCTTGTCTGCAAGGACTCAGCTGCA
GGTTTTATCGCCAACTTCTGTGCAGCGGCAGTTCTTCCCACTGATGAGGACGAGATAGG
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA
GAGGGGAACTTCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTTGGGATG
GGCCGCTCGGCTTACAATGAAGGGGACTATTATCATACGGTGTTGTGGATGGAGCAGGTGCT
AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCAACAGGTGCTGGACTACCTCA
GCTATGCTGTCTTCCAGTTGGGTGATCTGCAACGGTGCCCTGGAGCTCACCCGCCGCTGCTC
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA
GGAAGAGAGAGAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAAAGGCA
TCTATGAGAGGCTGTGGACTACCTGCCCTGAGAGGGATGTTACGAGAGCCTCTGTCTGTGGG
GAGGGTGTCAAACCTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACTATGGCAA
CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACA
TCGTCAAGTACTACGATGTCATGTCTGATGAGGAAATCAGAGAGGATCAAGGAGATGCAGAAA
CCTAAACTTGCACGAGCCACCGTTTCGTGATCCCAAGACAGGAGTCTCACTGTCTGCGCAGCTA
CCGGGTTTCCAAAAGCTCTGGCTAGAGGAAGATGATGACCTGTGTGGCCCCGAGTAAATC
GTCGATGACGATATCAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT
TATGGAGTGGGAGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG
CCTCAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGTATGTAAGCTG
GTGGTGCCACCGCTCTTCCCTGATCTGGGGCTGCAATTTGGCCCTAAGAAGGGTACAGCTGTG
TTCTGGTACAACCTCTTTCGGAGCGGGGAAGGTGACTACCGAACAAGCATGCTGCTGCTGCC
TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT
TGAGACCTTGTGGATCAACAGAAGTTGACTGAATCATCTTTCTGTCTTCCCTTCTCTGGTC
CTTCAGCCCCATGTCAACGTGACAGACACCTTTGTATGTTCCCTTTGTATGTTCTATCAGGCT
GATTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT
GTGACTGAAGTCCAGCCCTTCCATTACGCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCTAGCAAGGTGCCCTTGTAACCTCAGGTGTT
TTAGGTTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTGTTTTAT
GGCATTCTATCTATTGTGCTTTACCAAAAAATAAAATGTCCTACCAGAAAAA

FIGURE 116

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA
NYMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDVLQDSAAGFIANLSVQRQFFP
TDEDEIGAALKMRLQDTYRLDPGTISRGEPLGTTYQAMLSVDDCFGMGRSAYNEGDIYHTV
LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLSLDPSHERAGGNLR
YFEQLLEEREKTLTNQTEAELATPEGIYERPDVYLPEDVYESLCRGEVGLTPRRQKRLF
CRYHHGNRAPQLLIAPFKEEDEWDSPHIVRYDVMSDEEIERIKEIAKPKLARATVRDPKTG
VLTVAASYRVSKSSWLEEDDPVVARVNRMRQHITGLTVKTAELLQVANYGVGGQYEPHDFDS
RPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKGTAVFWYNLLRSGEVDYR
TRHAACPVLVGCKWVSNKWFHERGQEFRLPCGSTVD

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,
346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

FIGURE 117

GCAGTATTGAGTTTACTTCTCCTCTTTTGTAGTGAAGACAGACCATAATCCAGCTGTGAGTGAAATTGATTGT
 TTCAATTTATTACCGTTTGGCTGGGGGTAGTTCCGACACCTTCAAGTTGAAGAGCAGGCAGGAAGGATTGTGA
 AGACAGAGCAAACTCTTTGGGAGCTGTGTCCTGGAAGCCAGCGGCTTGTCTGTGCTTTGGCCTCATTGACCC
 CAGGTTCTCTGGTTAAACTGAAGGCTACTACTGGCTGGTGCCTCAATCCATTGATCCTTGAGGCTGTGCC
 CTGGGGCACCACCTCTGGCAGGCTTACCACATGCGACTGGCTCCCTGTGGCTCTGTCTGGGGCAGCGCTCT
 CCCCATCTTTAGGGCTGTCTCTGGGGTGCAGCCTGAGCCTCTGCGGGTTTCTGGATTCCAGGGGGAGGAGAAG
 ATCCTGTGTGTCAGGCTGTAGGGAGCGAGGAGGGCCACAGAATCCAGATTCCAGAGCTCGGCTAGACCAAAGTG
 ATGAAGACTTTCAAACCCCGATTGTCCCTACTACAGGAGCCCAACAGCCCTACAGAAGGTGCTCAGGACTC
 GGTAATCCAGACAGCTGGGCTCCGCTAGCGGTTGCTGGTGGCTGTCTGACTCCCGAGCTACACTGTCCA
 CTTTGGCCGTGTGCTGTGAACCTACGCTGGCCCATCACTTCCCTCGGTACTCTACTTCACTGGGAGCGGGGG
 CCGGCTCCAGCAGGATGCAGGTGGTGTCTCATGGGGATGAGCGGCCCGCTGGCTCATGTACAGACCTGC
 GCCACCTTCAACACACTTTTGGGGCCGACTACGACTGGTTCTCATCATGAGGATGACACATATGTGACGGCCC
 CCGGCTTGGCAGCCCTTGTCTGCCACCTCAGCATCAACCAAGACCTGTACTTATGGCCGGGAGAGGAGTTCATTG
 GCGCAGGCGAGCAGGCCCGGTACTGTCTATGGGGCTTTGGCTACCTGTTGTCCAGGAGTCTCTGTGCTTGTCTG
 GGCCACATCTGTGATGGCTGCCGAGGAGACATTTCTCAGTGGCCGCTCTGACGAGTGGCTTTGGAGCGCTGCCCTATTG
 ACTCTCTGGGGCTGGCTGTGTCTCAGCAGCAGGAGGCGAGTATCGCTCATTTGAACCTGGCCAAAATAGGG
 ACCCTGAGAAGGAAGGAGCTCGGCTTCTCTGAGTGCTTCCGCGTGACCCCTGTCTCCGAAGGTACCTCTCATGT
 ACCGGCTCCACAAACGCTTTCAGCGCTCTGGAGTTGGAGCGGCTTACAGTGAAATAGAACACTGCGAGGCTCAGA
 TCCGGAACCTGACCGTGCTGACCCCGAAGGGAGGCGAGGCTGAGCTGGCCCGTTGGGCTCCTGCTCCTTTTCA
 CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTTCACAGAGCAGCACACTTCTCTGTGTCAGATGGGG
 CTCCTCAAGTGCCCATCAGGGGGCTAGCAGGGCGAGCTGGGTGATGCGTTGGAGACTGCCCCGTGAGCAGCTCA
 ATCGGGCTATCAGCCCCGCTTGCCTTCCAGAAGCAGCGACTGTCTCAACGGCTATCGGCGCTTCGACCCAGCAC
 GGGCATGGAGTACACCTTGACCTGTCTGTGAATGTGTGACACAGCGTGGGCACCGGGGCCCTGGCTCGCA
 GGGTCAGCCTGCTCGGCCACTGAGCGGGTGGAAATCTTACTATGCCCTATGTCTACTGAGGCCACCCGAGTGC
 AGCTGTGTCTGCCACTCTGTGTGGCTGAAGCTGTCTGACGCCCGGCTTCTCTGAGGCGTTTGCAGCCAATGTCTC
 TGGAGCCACGAGAACTGATCTTGTCTCACCTGTGTGGTCTACGGGCACAGAGAAGGTGGCCGTGGAGCTCCAG
 ACCCAATTTCTGGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTGGCTCG
 CTGTGCGAGCAGAGGCCCTTCCAGGTGCGACTCATGAGCTGGTCTCGAAGAAGCACCTGTGAGACTCTCTCT
 TCTTCTTACCACCGTGTGGACAAGGCTGGGCCGGAAGTCTCAACCGCTGTGCGATGAATGCCATCTCTGGCT
 GGCAGGCTTCTTTCCAGTCCATTTCCAGGAGTTCAATCTCGCCTGTCAACACAGAGATACCCCCAGGGCCCC
 CGGGGGCTGGCTTGACCCCCCTCCCTCTGTGTGTGACCCCTCCCGGGGGCTTCTATAGGGGGAGAGATTG
 ACCGGCAGGCTTCTGCGGAGGGCTGCTTCTACAACGCTGACTACTGGCGGCCGAGCCTGGCTGGCAGGTGAAC
 TGGCAGGCCAGGAAGAGGAGGAGCCCTGGAGGGGCTGGAGGTGATGATGTTTCTCTCGGTTTCTCAGGGCTCC
 ACCTCTTTCCGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGACGCCACGGCTCAGTGAAG
 AACTCTACCACCGCTGCGCCTCAGCAACTGGAGGGGCTAGGGGGCGTGGCCAGCTGGCTATGGCTCTCTTTG
 AGCAGGAGCAGGCCAATAGCACTTAGCCCCCTGGGGGCCCTAACTCATTAACCTTTCTCTTTGTCTGGCTCAGCC
 CCAGGAAGGGCAGGCAGAGTGGTGGACAGATAGAGAATTGTGTCTGTATTTTAAATATGAAATGTTATTAA
 ACATGTCTTCTGCC

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

FIGURE 118

MRLSSLLALLRPALFLILGLSLGCSLLRVSWIQEGEGEDPCVEAVGERGGPQNPDSSRARLD
QSDQDFKPRIVPPYRDPNKPYKKVLRTRYIQTELGSRRERLLVAVLTSRATLSTLAVAVNRVT
AHHFPRLLYFTGQRGARAPAGMQVSVSHGDERPAWLMSETLRHLHTHFGADYDWFIMQDDTY
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLRLRLRPHLDGCRG
DILSARPDDEWLGRCLIDSLGVGCVSQHQQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPGEAGLSWPVGLPAPFTPHSRFEV
LGWDYFTEQHTFSCADGAPKCP LQGASRADVGDALETALEQLNRRYQPRLRFQKQRLNGYR
RFDPARMEYTLDDLLECVTQRGHRRALARRVSLRLPLSRVEILPMPYVTEATRVQLVLP LL
VAEAAAAAPAFLEAFAANVLEPREHALLTLLLVYGPREGGRGAPDPFLGVKAAAELERRYPG
TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLT TVWTRGPGEVLNRCRMNAISGWQAFPP
VHFQEFNPALSPQRSPPGPPGAGDPFSPPGADPSRGAPIGGRFDRQASAEGCFYNADYLA
RARLAGELAGQEEFEALGLEVMDVFLRFSGHLHFRAVEPGLVQKFSRLDCSPRLSEELYHR
CRLSNLEGLGGRAQLAMALFEQEQA NST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTTTCCTGTCCCCAAGCC
GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG
AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA
TGCACCACCATGAGCATCATCACCTACAAGCTCCTTAACAAAGAAGATATCTTGAAAAATTTCA
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTTCGAGTATACTGTATTATCCTTGTA AAAACC
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACTGTGACAAAGCAG
AGTTCTTCAGTCTGAAAAATGTTAAAGTGTTCAGTCAATTAATATGGACACAAATGACATG
TGGTTAATGATGAGAAAAGCTTACAAATACGCCCTTTGATAAGTATAGAGACCAATACAACGTG
GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTGT TAA
AAAAGGATCCATCACAGCCTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT
GTGGGTATGGAAGGAGGAATGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT
CAATATCCCAGAAAAGTGTCTGAACAGGGAGGGATGATTGGAAGATATCTGAAGATAAAC
AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA
AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC
CAACCAGGTAGTAGAAGGCTGTTGTT CAGATATGGCTGTTACTTTTAATGGACTGACTCCAA
ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATTGGGCATATTTTCAAT
GATGCATTGGTTTTCTTACCTC CAAATGGTTCTGACAATGACTGGAAGTGGTAGAAAAGCG
TGAATATGATCTTTGTATAGGACGTGTGTTGT CATTATTTGTAGTAGTAAC TACATATCCAA
TACAGCTGTATGTTCTTTTCTTTTCTAATTTGGTGGCACTGGTATAACCAACACATTAAAG
TCAGTAGTACATTTTAAATGAGGGTGGTTTTTTCTTTTAAACACATGAACATTGTAATG
TGTGGAAAGAAGTGT TTAAGAATAATAATTTGCAAATAAACTATTAATAAATATTATAT
GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGCACATATTTTGTCTGATTGGTT
AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT
TGTGATTAAAGTAAACCTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT
AAGCCTCCCCAAGTCCAATGGATTGCGCTTCTCAAATGTACAACCTAAGCAACTAAAGAAA
ATTAAAGTGAAAGTTGAAAAAT

FIGURE 120

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISDERME
LSKSFVRVYCIILVKPKDVSLWAAVKETWTKHCDKAFFSSSENVKVFESINMDTNDMWLMMRK
AYKYAFDKYRDQYNWFFLARPTTFATIENTLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGG
IVLSVESMKRLNSLLNIPEKCPEQGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFN
TSVGLSLIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFPGHIFNDALVFL
PPNGSDND

Signal sequence:

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

[illegible][illegible]

FIGURE 122

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN
FTELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSSAMGAHLVVINSQEEQ
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCAITMRDSS
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

Signal sequence:

amino acids 1-42

N-glycosylation site.

amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.

amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38

100
90
80
70
60
50
40
30
20
10
0

[illegible]

GGGA CTCAGAGCCGGCCGCT GCGTCGCGCTGGCCCTCAGCAACCTCGACATGGCGCGTGAGCGGCCCCCGGAC
TCGGCTCTTCGGCTCGGCTCGCTGCACTTCTTCCTGCTGCTGCTTTTCAGGGCGTGCCTGTAGTGGGGCTGTAAATCT
TCAAACTCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAGTGTGGAAATGCTTGTGCATTAACCGATTTCG
AGACAAGTGAACCCAGAGTCCGATGAGGAAATTAACAGATGAACAAACCATATGTGTTTGTGACAAACAAA
TTCAGGGAGACTTGGCGGTGCTGCAGAAATCTGGGAGAGATCCCTGAAGATCTTGGATGTGACACAGGAGAG
ACTCAGACCCCTTTATCGCTGTGAGGCTGTGCTCGAAATGACCGCAGGAAGAAATGATAGATGTGATCGAGTTAA
CTGTGCAAGTGAAGAGTGACACCCCTGTCTGTAGATGTCCAGAGCTGTACAGTGAAGCCAGATGGACAACTAC
ACTGCGCAGGAGATGGAGGCGACCCCGCCCTCACTACAGCTGTGATGTGCAATGATGTACCATGCCACCGAT
CCAGAGCGAATCCAGATTTTGGCAATCTCTTTCTCACTTAAACTGTGAAGAGGCATTTGGTGTGCTACTGCTG
TTCACAAAGCAGCATCTTGGGAGTACTGATTCGATTCGTTTCCAAATGACGAGGCTCAGCCAGGTGTGAGGAGCAGG
AGATGGAAGTCTATGACCTGCAACTTGGCGAATTAATGGGGGGGCTTGTGTTGCTGCTGCTACTGGCCCTGA
TCAGCTTGGCATCTGTCTGTGCATACAGAGTGCTGCTACTTCACTCAACCAATAAACAGATGGAGAAGTTTACAAGA
ACCCAGGGAAACAGATGGAGTTTAACATACTCGCATCGCATCAGGAGGAGGCGCTCTAGACAAGAGTCTATGTTT
TGATCTGGACCGCGGCTGTGGCTCAGAGCGCACAGAGCGACAGTGCACATACATCTTGCTGTACATACCTCTGTGCA
GGCAGCGAGGCTGATGCACTCGACAGCGCTAGACATCACTCAGAGCTTTTCTGTTTGGCCAAAGTGTAACCA
CTACTCTTCTTACTTAAACAGGCGCATGAATAGGAGAAATTTCTCAGAGTGAACCCGGGTGAACATAAACACCAA
GGAAGCGAAATGGGTGGCTTCACTGATTTGGTTCCTAACTGTTTGTGGCTGCTTCCGATAGTATTAGG
GTGATCTTAAAGAGTTTGTCTGCACTTAAACCGCGCTGGGCGCTGTGAAGCCAGCATGTCTCCACGTGTGCTT
CAGCAGCCACAGCAGACCATGTGATGATGGCAGTGCTGCGGCACAGCAGCAGCGCATCCCGCGGGAAACCA
GAAAGGCTCTTCTTACAAGCAGCCCTTACTGATCTCGGCCCAAGACACCCAGCGCTTCTCTTTAAAGCTCTGCT
TGATCGGTGTGCAGTGTCCATTGTAGGAAGCTTTTGGATCAGCAATTTGTAAAAACAACAAATCAGGAAG
GTAAATTTGGTGTGGAAGAGGGATCTGTGCTGAGGAACCTGTCTGTCCACAGCGGTTCAGATTTTAAGGAA
ACCTTGCTCTTAGGCTGTGCTGAATGTGACTGAATATGCTTTTCTATGGGCTTGTGTTATTTATAAAAT
TACATCTAAATTTTCTTAAGATGTATTTTGATATTGAAAGGAAATTTCTATTAAATCTGAATATATTGT
CATCAATGTAAATAACCTATTTTAAAAAGATTCAACTTAAGGTGAAGTTTCAAGTCACTAGTGTTTAAAT
TGGAAATATCAATAATTAAGATTAATTTTCCACAGGAATCCTCATGGAAGTTTACTGTGATGTCTCTTTAACT
CACACAGGTTTATGCTTTTTCACAGGGAATCTATGCTTCACATCAGACCATGATGCTTGAAGAAACCTT
TAAAAATCCAGTTAGCAATGTGAAATCAGTTTGCATCTCTCAAGAAAGAACTCTCAGGTAGCTTTGAATC
GCTCTCTCTCAGATATGATGAGCATCTGTGACCCAGGCGCACCCAGAGAGCCCTCAGATGTACATACAGATG
CCAGTCTGCTCTGGGTTGGCGCAGGCGCCCGCTCAGTCACTGTTGCTCTGCTGTCTGCGCAGAGGCCCT
GCATCTCTGGGCGCTGAGCGTGGCTGTGCTCCAGTGAGCTTTACTCACTGGGCCCTTGCTCTCAGACACAGC
TTCAGGTGGCACTCAGGGAACCTGTGCTGTCTTCAATGAGCGCCAGCTTTGGGCTCTGTAAACAGACCTCT
TTTGTGTTATGATGGTCTCAAAAATAGGCGGCCCAATGCTATTTTTTTTTTAAAGTTTGTTTTAAATTTATTTT
AAGATTGTCATAGGCGAAAGGCAATGTGCAAAATCAAGTCTGTCAAGTACAATAACATTTTAAAAAGAAATGAT
CCCACTGTTTCTCTTGTGCACAGAGAAGACCCAGAGCGCACAGGCTCTGTGCAATTTCAAAACAACCATGAT
GGAGTGGCGGCGCTCAGCCCTTTTAAAGRAAGTCAGTGTGAGCAGCAGGTGTGAAGAGGCTCGGGGGAGGAAG
TGAAACCGCTGAATCAAAAGCACTTTCTAATTTTGACTTTAAATTTTTCATCGCGCAGAGACATCTGCTCCAAAT
TGTGTGGGGGACATTAGCAACTCATTCAAGAGCGCTGTGTTCTCAAGCAGAGTGTGTTCTCAGCCTCACAATGCCCT
CGCGTGTGATCAGGACTGAAGTGTGCTGAAAGCAAGAGCTGTGAGAAGGAACATCCCACTGTGCTCGTGA
GATGTGCTCTCACTACCTGTTCTTTCAGCTTCCAGTCTTGTGGTTTTTTATCTTTAGAGCTTTTTTTTT
AATTCATCATAGACTGATGTGTGACTTTTGTGTTTGTGATTTGTGAAACATTTTGGCAGAGCGCGCTTGGCAGAGGCA
GAAATTTGCTCAGCAGCTGCTCAGTCTCCCTGTTCTGCTGCTCATGGCATCTGGATGTCTAGATCAGAGTTT
CTCTCATCATGTCACCTTGTAGAGAGGGATGAGTCCCCACCTCAGGCTTGGGAGTTCAGCCTCAGCGCTCCAT
CTTTGTGTGCTATAGTATGAGGTAGCTTCTTGTCCCTCTCTTATACCTTAAAAAACCCTTACATAGTGTGCT
TGGGAACAGAGTCTGAAAAAGTAGAGAGAGTGAAGATGAGTCTGGGAAGTACGCTGCTTATAGATGAGACTAGA
CGGAAAGGAATACTCGTGTATTTTAAAGATGAATGTGACTCAGATCAGGCGCTCAGGCGCTGATGTTCT
GCTTTGTGATGATGTGTGCTGTACACAGTGCTACAGACTGTGATCACTAACACACCGTAAATTTGCGATTTGTTTAACT
CTCAATTTTAAAGAGCTTCAAAAACACCA

FIGURE 124

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624
><subunit 1 of 1, 310 aa, 1 stop
><MW: 35020, pI: 7.90, NX(S/T): 3
MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTFVVQEFESVELSCIITDSQTS
PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLKINNVTRRDSALYRCEVVARNDRK
EIDEIVIELTVQVKPVTVPVCRVPKAVPVGKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRA
NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDLNLIGGIIG
VLVVLAVLALITLGICCAYYRRGYFINNKQDGESYKNPGKPDGVNYYIRTDEEGDFRHKSSFVI
```

Important features of the protein:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267